

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/04, C12N 15/11, C12P 21/02	A1	(11) International Publication Number: WO 97/45439 (43) International Publication Date: 4 December 1997 (04.12.97)
(21) International Application Number: PCT/US97/09460 (22) International Filing Date: 30 May 1997 (30.05.97) (30) Priority Data: 08/657,620 30 May 1996 (30.05.96) US (60) Parent Application or Grant (63) Related by Continuation US 08/657,620 (CIP) Filed on 30 May 1996 (30.05.96) (71) Applicant (for all designated States except US): THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; West 116th Street and Broadway, New York, NY 10027 (US). (72) Inventor; and (75) Inventor/Applicant (for US only): STURLEY, Stephen, L. [GB/US]; 410 Riverside Drive #34, New York, NY 10025 (US). (74) Agent: WHITE, John, P.; Cooper & Dunham LLP, 1185 Avenue of the Americas, New York, NY 10036 (US).		(81) Designated States: AU, CA, JP, MX, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: DNA ENCODING ACYLCOENZYME A: CHOLESTEROL ACYLTRANSFERASE AND USES THEREOF		
(57) Abstract This invention provides an isolated nucleic acid which encodes an acylcoenzyme A: cholesterol acyltransferase II or III. Specifically, this invention provides an isolated nucleic acid which encodes a human wildtype acylcoenzyme A: cholesterol acyltransferase II or III. This invention also provides various methods for inhibiting wildtype acylcoenzyme A: cholesterol acyltransferase II or III in a subject. This invention also provides a method for identifying a chemical compound which is capable of inhibiting acylcoenzyme A: cholesterol acyltransferase II or III and a pharmaceutical composition comprising of the chemical compound identified by the above-described method. This invention also provides a method of treating a subject who has atherosclerosis or hyperlipidemia.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

**DNA ENCODING ACYLCOENZYME A: CHOLESTEROL
ACYLTRANSFERASE AND USES THEREOF**

5

This application is a Continuation-In-Part of U.S. Serial No. 08/657,620, filed May 30, 1996, the content of which is incorporated by reference into this application.

10 Throughout this application, various publications are referenced by Arabic numerals. Full citations for these publications may be found listed at the end of the specification. The disclosures of these publications in their entireties are hereby incorporated by reference
15 into this application in order to more fully describe the state of the art as known to those skilled therein.

Background of the Invention

Cholesterol or related sterols, required for the
20 viability of eukaryotic cells, exist in the free form or as esters conjugated to fatty acids. The concentration of free sterol determines the fluidity of eukaryotic cell membranes, whereas esterified sterols cannot participate in membrane assembly. The esterification of
25 intracellular sterol, mediated in mammals by the membrane-bound enzyme, acylcoenzyme A: cholesterol acyltransferase, is thus a critical homeostatic determinant of membrane function (1, 2). For example, cholesterol depletion of the rough endoplasmic reticulum
30 (ER) relative to the smooth ER (3), may modulate protein translocation or membrane-associated transcriptional activators such as the Sterol Response Element Binding proteins (SREBP, 4). In addition, production of cholesterol ester (CE) by acylcoenzyme A: cholesterol
35 acyltransferase in the rough ER may influence the transport of sterol between intracellular pools. Similar

-2-

esterification activities have been observed in other eukaryotes such as plants and yeasts (5).

5 Elevations in acylcoenzyme A: cholesterol acyltransferase activity perturb several pathways that contribute to hyperlipidemia and atherosclerosis. Sterol esterification modifies the activity of the low density lipoprotein (LDL) receptor and alters serum lipoprotein composition to be pro-atherogenic (6, 7). It may also be
10 a rate limiting step in intestinal sterol absorption (8). Furthermore, CE deposition in the arterial wall is an important initial step in atherogenesis (9). The understanding of the acylcoenzyme A: cholesterol acyltransferase reaction has been hampered by the
15 difficulty of biochemical purification and by a poor grasp of the relevant genetic determinants. A human acylcoenzyme A: cholesterol acyltransferase I gene from macrophages was identified by complementation of Chinese Hamster Ovary cell lines deficient in acylcoenzyme A:
20 cholesterol acyltransferase activity (10) and was functionally expressed in insect cells devoid of endogenous activity (11).

25

-3-

Summary of the Invention

This invention provides an isolated nucleic acid which encodes an acylcoenzyme A: cholesterol acyltransferase II
5 or an acylcoenzyme A: cholesterol acyltransferase III.

This invention also provides a vector which includes the isolated nucleic acid which encodes an acylcoenzyme A: cholesterol acyltransferase II or an acylcoenzyme A: cholesterol acyltransferase III and a host vector system
10 which includes a vector.

This invention also provides a method of producing a polypeptide which comprises growing such host vector system of claim 14 under suitable conditions permitting
15 production of the polypeptide and recovering the polypeptide so produced. This invention also provides a purified wildtype acylcoenzyme A: cholesterol acyltransferase II or an acylcoenzyme A: cholesterol
20 acyltransferase III.

This invention also provides an oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within a
25 nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or an acylcoenzyme A: cholesterol acyltransferase III without hybridizing to a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II or an acylcoenzyme A:
30 cholesterol acyltransferase III. This invention also provides an oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within the nucleic acid which encodes a mutant acylcoenzyme A: cholesterol
35 acyltransferase II or an acylcoenzyme A: cholesterol

-4-

acyltransferase III without hybridizing to a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or an acylcoenzyme A: cholesterol acyltransferase III.

5

This invention also provides a method for determining whether a subject known to have an imbalance in sterol levels has the imbalance due to a defect in esterification of sterol and for treating a subject who
10 has an imbalance in sterol levels due to a defect in esterification of sterol.

This invention also provides methods for inhibiting wildtype acylcoenzyme A: cholesterol acyltransferase II
15 or an acylcoenzyme A: cholesterol acyltransferase III in a subject.

This invention also provides a method for identifying a chemical compound which is capable of inhibiting
20 acylcoenzyme A: cholesterol acyltransferase II or an acylcoenzyme A: cholesterol acyltransferase III in a subject and a pharmaceutical composition comprising of the chemical compound so identified.

25 This invention also provides a transgenic, nonhuman mammal comprising the isolated nucleic acid which encodes acylcoenzyme A: cholesterol acyltransferase II or an acylcoenzyme A: cholesterol acyltransferase III.

30

Brief Description of the Figures

Abbreviations: The amino acid residues are abbreviated as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. CON: consensus sequence.

Figures 1A and 1B. Protein sequence alignments predicted from candidate genes for the human acylcoenzyme A: cholesterol acyltransferase gene I, the yeast homologs, *acylcoenzyme A: cholesterol acyltransferase-related enzyme I* and *acylcoenzyme A: cholesterol acyltransferase-related enzyme II*, and a consensus sequence of all three sequences.

Identical residues between all the sequences are in bold face. Residues of the candidate leucine zipper heptad motif are italicized. Potential transmembrane domains were identified at residues 132 to 155 and 460 to 483; 186 to 202 and 406 to 421; and 215 to 231 and 439 to 451, for human acylcoenzyme A: cholesterol acyltransferase (Sequence I.D. No.: 2), *acylcoenzyme A: cholesterol acyltransferase-related enzyme I* (Sequence I.D. No.: 4) and *acylcoenzyme A: cholesterol acyltransferase-related enzyme II* (Sequence I.D. No.: 6), respectively. The firefly luciferase signature sequences identified in human acylcoenzyme A: cholesterol acyltransferase I (10) were not conserved in the yeast genes. CON (Sequence I.D. No.: 13) denotes the consensus sequence between the sequences of human acylcoenzyme A: cholesterol acyltransferase, *acylcoenzyme A:*

-6-

- cholesterol acyltransferase-related enzyme I and acylcoenzyme A: cholesterol acyltransferase-related enzyme II. R07932 denotes the partial sequence of another human acylcoenzyme A: cholesterol acyltransferase candidate cDNA (residues 500 to 600) (Sequence I.D. No.: 14). The asterisks indicate the residues in R07932 identical to those of the other sequences.
- 5
- 10 1A. Alignment of amino acid residues 1-362 of acylcoenzyme A: cholesterol acyltransferase-related enzyme I and the identical residues in acylcoenzyme A: cholesterol acyltransferase-related enzyme II, human acylcoenzyme A: cholesterol acyltransferase and CON.
- 15
- 1B. Alignment of amino acid residues 363-611 of acylcoenzyme A: cholesterol acyltransferase-related enzyme I and the identical residues in acylcoenzyme A: cholesterol acyltransferase-related enzyme II, human acylcoenzyme A: cholesterol acyltransferase and CON.
- 20

25 **Figures 2A, 2B, 2C, 2D and 2E. Construction and analysis of acylcoenzyme A: cholesterol acyltransferase genes and deletion mutants.**

- 2A. The *are1* DNA deletion. The schematic depicts a fragment from yeast chromosome III in plasmid pH3(34). Strategic restriction endonucleases are indicated (H, Hind III; B, Bam HI).
- 30 2B. The autoradiogram depicts Bam HI digested DNA from wild-type or disrupted diploid strains probed with the 2993-bp Bam-HI fragment. This produced a fragment corresponding to the wild-type acylcoenzyme A: cholesterol

-7-

acyltransferase-related enzyme I locus and a 1984-bp fragment characterizing the *are1Δ* NA allele. The diploid is heterozygous for the acylcoenzyme A: cholesterol acyltransferase-related enzyme I deletion.

5
2C. Reduced stringency hybridization of yeast genomic DNA with acylcoenzyme A: cholesterol acyltransferase-related enzyme I coding sequences. Genomic DNA from wild-type or
10 *ARE1/are1Δ*NA diploids were reprobbed with an Nhe I-Avr II fragment corresponding to the acylcoenzyme A: cholesterol acyltransferase-related enzyme I open reading frame ("ORF"). Hybridizations and washes were performed at
15 60°C in the absence of formamide.

2D. The *are2Δ* deletion. In step 1, PCR amplifying oligonucleotides, KO-5' and KO-3' and a *LEU2* template were used to produce the selectable yeast gene flanked at the 5' and 3' ends by
20 acylcoenzyme A: cholesterol acyltransferase-related enzyme II. In step 2, this was used to direct homologous recombination at acylcoenzyme A: cholesterol acyltransferase-related enzyme II by transformation of a diploid strain and selection for leucine protrophy. In step 3,
25 integrants to acylcoenzyme A: cholesterol acyltransferase-related enzyme II were identified by a PCR reaction using oligonucleotides flanking *ARE2* (*are2*-5' and *are2*-3') and a 3' amplimer within *LEU2* (L2-3').
30

2E. A 999-bp fragment identifies *are2Δ*, as shown in the ethidium bromide stained agarose gel. The wild-type fragment (2206-bp) is also produced

-8-

in the same reaction. Leucine prototrophic transformants with deletions of acylcoenzyme A: cholesterol acyltransferase-related enzyme II were obtained at a frequency of ~2%. M indicates the 50-2,000-bp ladder markers (Bio-Rad Laboratories).

Figures 3A and 3B. Fluorescent staining of triglyceride and sterol ester.

The cells were grown in YEPD to stationary phase, washed with deionized H₂O, and incubated with 1 µg/ml Nile Red (1 mg/ml in acetone). Fluorescent images were obtained with a BioRad MRC600 laser scanning confocal microscope (BioRad Microscience, Hercules, CA) on an inverted Zeiss Atiovert microscope (Zeiss, OberKochem, Germany) using 63X (NA1.4) Zeiss Plan-apo infinity corrected objective. Samples were illuminated with the 488nm line from an argon ion laser and the fluorescence was visualized with a 540nm dichroic mirror and 550nm long-pass emission filter. Staining of the cytoplasmic lipid droplets was sensitive to treatment with isopropanol, proving them to be lipid in nature.

3A. Wild-type cells.

3B. *are1ΔNAare2Δ* double mutant cells.

Figures 4A, 4B, 4C and 4D. Neutral lipid and sterol biosynthesis in ARE deletion mutants.

Strain genotypes are as described in the text; dpm/mg dry weight: disintegrations per minute per milligram of dry weight of cells.

4A. Triglyceride biosynthesis. Total lipids were

-9-

extracted from cells grown in media containing ^3H -oleate and analyzed by thin-layer chromatography.

4B. Sterol ester biosynthesis. Total lipids were extracted from cells grown in media containing ^3H -oleate and analyzed by thin-layer chromatography.

4C. Sterol ester biosynthesis in wild-type and mutant cells transformed with vector control (black box) or acylcoenzyme A: cholesterol acyltransferase-related enzyme I over-expression plasmids, YEp3-16 (increased copy number, shaded box) and pADH5-36 (transcription from the ADH promoter, open boxes). Cells were grown in selective media to maintain the acylcoenzyme A: cholesterol acyltransferase-related enzyme I expression plasmids. Lipids were labeled, extracted and analyzed as above.

4D. Sterol biosynthesis in acylcoenzyme A: cholesterol acyltransferase-related enzyme deletion mutants. Lipids were labeled in synthetic complete media containing $[1-^{14}\text{C}]$ acetate, saponified and extracted with hexane and subjected to thin layer chromatography analysis. The data is representative of three separate experiments and expressed as the ratio of incorporation into sterols to incorporation into fatty acids.

Figures 5A, 5B, 5C, 5D, 5E and 5F. The nucleic acid and amino acid or predicted amino acid sequences.

5A-1 - 5A-3.

The nucleic acid sequence of human acylcoenzyme A: cholesterol acyltransferase I designated

-10-

Sequence ID No.: 1. The amino acid sequence of human acylcoenzyme A: cholesterol acyltransferase I designated Sequence ID No.: 2.

- 5 5A-1. Nucleic acid sequence of human acylcoenzyme A: cholesterol acyltransferase I from nucleic acid bases 1-1624. Amino acid sequence of human acylcoenzyme A: cholesterol acyltransferase I from amino acid residues 1-76.
- 10 5A-2. Nucleic acid sequence of human acylcoenzyme A: cholesterol acyltransferase I from nucleic acid bases 1625-2524. Amino acid sequence of human acylcoenzyme A: cholesterol acyltransferase I from amino acid residues 77-376.
- 15 5A-3. Nucleic acid sequence of human acylcoenzyme A: cholesterol acyltransferase I from nucleic acid bases 2525-3649. Amino acid sequence of human acylcoenzyme A: cholesterol acyltransferase I from amino acid residues 377-551.
- 20
- 25

5B-1 - 5B-3.

30 The nucleic acid sequence of yeast acylcoenzyme A: cholesterol acyltransferase-related enzyme I designated Sequence ID No.: 3. The amino acid sequence of yeast acylcoenzyme A: cholesterol acyltransferase-related enzyme I designated Sequence ID No.: 4.

- 35 5B-1. Nucleic acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme I from nucleic acid

-11-

bases 1-1289. Amino acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme I from amino acid residues 1-209.

5

5B-2.

Nucleic acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme I from nucleic acid bases 1290-2114. Amino acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme I from amino acid residues 210-484.

10

5B-3.

Nucleic acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme I from nucleic acid bases 2115-2601. Amino acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme I from amino acid residues 485-611.

15

5C-1 - 5C-3.

20

The nucleic acid sequence of yeast acylcoenzyme A: cholesterol acyltransferase-related enzyme II designated Sequence ID No.: 5. The amino acid sequence of yeast acylcoenzyme A: cholesterol acyltransferase-related enzyme II designated Sequence ID No.: 6.

25

5C-1.

Nucleic acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme II from nucleic acid bases 1-1061. Amino acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme II from amino acid residues 1-238.

30

5C-2.

Nucleic acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme II from nucleic acid

35

-12-

bases 1062-1961. Amino acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme II from amino acid residues 239-538.

5 5C-3. Nucleic acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme II from nucleic acid bases 1962-2421. Amino acid sequence of acylcoenzyme A: cholesterol acyltransferase-related enzyme II
10 from amino acid residues 539-643.

5D. The nucleic acid sequence of mouse acylcoenzyme A: cholesterol acyltransferase II designated
15 Sequence ID No.: 11. The amino acid sequence of mouse acylcoenzyme A: cholesterol acyltransferase II designated Sequence ID No.:
12.

20 **Figure 6A. A restriction map of the expression vector YepAB-ACAT2.**

Figure 6B and 6C. Expression of human macrophage ACAT in pRS426GP.

25 6B. The ACAT open reading frame was inserted at the NotI and SacI sites, downstream of the promoter of the *GAL1/10* gene (*GAL1/10p*) as described in the text to produce pRS426-ACAT. *URA3* and *Amp^r* denote
30 selectable markers for yeast and *E. coli* respectively. The yeast and bacterial origins of replication (*2μm* and *ori*, respectively) are indicated.

6C. Immunoblot of human ACAT in protein

-13-

5 extracts from cells transformed with pRS426-ACAT. Double mutant cells (*are1⁻ are2⁻*), transformed with pRS426-ACAT (hACAT) or with pRS426GP (vector), were induced by growth in galactose. Proteins were analyzed by immunoblotting. Equivalent amounts of protein extracts from mouse adrenal cells were loaded for comparison. Molecular weight reference markers (BioRad) are indicated (M). The arrow indicates the position of the DM10 immunoreactive product in extracts from murine adrenals. The expressed form of hACAT in yeast is of coincident mobility.

10
15 **Figures 7A and 7B. Multiple human tissue Northern analysis of poly (A)+ RNAs probed with ³²P-labeled cDNA C1.**

20 7A. Tissue specific expression of wildtype human acylcoenzyme A: cholesterol acyltransferase II using a wildtype acylcoenzyme A: cholesterol acyltransferase II specific probe.

25 7B. Tissue specific expression of wildtype human acylcoenzyme A: cholesterol acyltransferase I using a wildtype acylcoenzyme A: cholesterol acyltransferase I specific probe.

Figure 8A, 8B, 8C and 8D. Tissue specific expression of ARGP1 and hACAT.

30 8A and 8B. Multiple tissue Northern (Clontech) with indicated samples were probed with an ARGP1 specific probe as described in the text.

8C and 8D. The same blots were also analyzed using a hACAT specific probe. The first panel is

-14-

identical to that published by Chang et al (8). The second panel is the same blot as in A and B, probed with the ACAT cDNA 1600 bp probe.

5

Figure 9. Fetal Tissue specific expression of AGRP2.

Multiple tissue Northern blots of fetal tissue (Clontech) with indicated samples, were probed with an AGRP2 specific probe as described in the text.

10

Figure 10. Cultured cell expression of AGRP1.

RNA samples from HepG2 and CV1 were reverse transcribed and PCR amplified as described in the text. P indicate a plasmid template control. The blank lanes represent water or no RT controls.

15

Figure 11. Sequence comparison of human ACAT and AGRP1

Figure 12. Sequence comparison of human ACAT and AGRP2

20

Figure 13. Phylogenetic Comparisons of ACAT like molecules.

The sequences shown were identified in genome databases and aligned based on protein sequence using GCG Inc software (pileup). They were subsequently arranged to their sequence conservation to determine approximate evolutionary relatedness.

25

Figure 14. Conserved motifs in ACAT related gene products.

30

Figure 15A and 15B. Nucleotide and predicted protein sequence of AGRP1.

-15-

Figure 16. Nucleotide and predicted protein sequence of
ARGP2.

-16-

Detailed Description of the Invention

Throughout this application, references to specific nucleotides are to nucleotides present on the coding strand of the nucleic acid. The following standard abbreviations are used throughout the specification to indicate specific nucleotides:

	C=cytosine	A=adenosine
10	T=thymidine	G=guanosine

A "gene" means a nucleic acid molecule, the sequence of which includes all the information required for the normal regulated production of a particular protein, including the structural coding sequence, promoters and enhancers.

The nucleic acids or oligonucleotides of the subject invention also include nucleic acids or oligonucleotides coding for polypeptide analogs, fragments or derivatives which differ from naturally-occurring forms in terms of the identity or location of one or more amino acid residues (deletion analogs containing less than all of the residues specified for the protein, substitution analogs wherein one or more residues specified are replaced by other residues and addition analogs where in one or more amino acid residues is added to a terminal or medial portion of the polypeptides) and which share some or all properties of naturally-occurring forms. These nucleic acids or oligonucleotides include: the incorporation of codons "preferred" for expression by selected non-mammalian hosts; the provision of sites for cleavage by restriction endonuclease enzymes; and the provision of additional initial, terminal or intermediate DNA sequences that facilitate construction of readily

-17-

expressed vectors.

5 The nucleic acids and oligonucleotides described and
claimed herein are useful for the information which they
provide concerning the amino acid sequence of the
polypeptide and as products for the large scale synthesis
of the polypeptide by a variety of recombinant
techniques. The molecule is useful for generating new
cloning and expression vectors, transformed and
10 transfected prokaryotic and eukaryotic host cells, and
new and useful methods for cultured growth of such host
cells capable of expression of the polypeptide and
related products.

15 An isolated nucleic acid which encodes an acylcoenzyme A:
cholesterol acyltransferase II. This isolated nucleic
acid may be DNA or RNA, specifically cDNA or genomic DNA.
Specifically, the isolated nucleic acid has the sequence
designated Seq. I.D. No.: 7. The isolated nucleic acid
20 encodes a human wildtype acylcoenzyme A: cholesterol
acyltransferase II having substantially the same amino
acid sequence as the sequence designated Seq. I.D. No.:
8. Specifically the isolated nucleic acid has the
sequence designated Seq. I.D. No.: 11. The isolated
25 nucleic acid encodes a mouse wildtype acylcoenzyme A:
cholesterol acyltransferase II having substantially the
same amino acid sequence as the sequence designated Seq.
I.D. No.: 12. Further, the isolated nucleic acid of
encodes a mutant acylcoenzyme A: cholesterol
30 acyltransferase II.

An isolated nucleic acid which encodes an acylcoenzyme A:
cholesterol acyltransferase III. This isolated nucleic
acid may be DNA or RNA, specifically cDNA or genomic DNA.
35 Specifically, the isolated nucleic acid has the sequence

as set forth in Fig. 16. The isolated nucleic acid encodes a human wildtype acylcoenzyme A: cholesterol acyltransferase III having substantially the same amino acid sequence as set forth in Fig. 16. Further, the
5 isolated nucleic acid of encodes a mutant acylcoenzyme A: cholesterol acyltransferase III.

As used in this application, "acylcoenzyme A: cholesterol acyltransferase III" means and includes any polypeptide
10 having acylcoenzyme A: cholesterol acyltransferase III activity and having an amino acid sequence homologous to the amino acid sequence of human acylcoenzyme A: cholesterol acyltransferase II (the sequence of which is set forth in Fig. 15). Thus, this term includes any such
15 polypeptide whether naturally occurring and obtained by purification from natural sources or non-naturally occurring and obtained synthetically, e.g. by recombinant DNA procedures. Moreover, the term includes any such polypeptide whether its sequence is substantially the
20 same as, or identical to the sequence of any mammalian homolog of the human polypeptide, e.g. murine, bovine, porcine, etc. homologs. Additionally, the term includes mutants or other variants of any of the foregoing which retain at least some of the enzymatic activity of
25 nonmutants or nonvariants.

As used in this application, "acylcoenzyme A: cholesterol acyltransferase II" means and includes any polypeptide
30 having acylcoenzyme A: cholesterol acyltransferase II activity and having an amino acid sequence homologous to the amino acid sequence of human acylcoenzyme A: cholesterol acyltransferase III (the sequence of which is set forth in Fig. 16). Thus, this term includes any such polypeptide whether naturally occurring and obtained by
35 purification from natural sources or non-naturally

-19-

occurring and obtained synthetically, e.g. by recombinant DNA procedures. Moreover, the term includes any such polypeptide whether its sequence is substantially the same as, or identical to the sequence of any mammalian homolog of the human polypeptide, e.g. murine, bovine, porcine, etc. homologs. Additionally, the term includes mutants or other variants of any of the foregoing which retain at least some of the enzymatic activity of nonmutants or nonvariants.

10

The invention also encompasses DNAs and cDNAs which encode amino acid sequences which differ from those of acylcoenzyme A: cholesterol acyltransferase II, but which do not produce phenotypic changes.

15

The invention also encompasses DNAs and cDNAs which encode amino acid sequences which differ from those of acylcoenzyme A: cholesterol acyltransferase III, but which do not produce phenotypic changes.

20

The nucleic acid of the subject invention also include nucleic acids that encode for polypeptide analogs, fragments or derivatives which differ from naturally-occurring forms in terms of the identity or location of one or more amino acid residues (including deletion analogs containing less than all of the residues specified for the protein, substitution analogs wherein one or more residues specified are replaced by other residues and addition analogs wherein one or more amino acid residues is added to a terminal or medial portion of the polypeptides) and which share some or all properties of the naturally-occurring forms.

25

30

35

The polypeptide of the subject invention also includes analogs, fragments or derivatives which differ from

-20-

naturally-occurring forms, but having acylcoenzyme A: cholesterol acyltransferase activity.

5 This invention also provides a vector comprising an isolated nucleic acid encoding acylcoenzyme A: cholesterol acyltransferase II or III. The isolated nucleic acid of the vectors is operatively linked to a promoter of RNA transcription which maybe, or is identical to, a bacterial, yeast, insect or mammalian
10 promoter. The vector may be a plasmid, cosmid, yeast artificial chromosome (YAC), bacteriophage or eukaryotic viral DNA. Specifically, this invention provides a vector designated YepAB-ACAT2 (Figure 6).

15 Further other numerous vector backbones known in the art as useful for expressing proteins may be employed. Such vectors include but are not limited to: adenovirus, simian virus 40 (SV40), cytomegalovirus (CMV), mouse mammary tumor virus (MMTV), Moloney murine leukemia
20 virus, murine sarcoma virus, and Rous sarcoma virus, DNA delivery systems, i.e liposomes, and expression plasmid delivery systems.

25 This invention also provides a vector system for the production of a polypeptide which comprises the vector in a suitable host. Suitable host includes a cell which includes, but is not limited, prokaryotic or eukaryotic cells, e.g. bacterial cells (including gram positive cells), yeast cells, fungal cells, insect cells and
30 animal cells.

Suitable animal cells include, but are not limited to, HeLa cells, Cos cells, CV1 cells and various primary mammalian cells. Numerous mammalian cells may be used as
35 hosts, including, but not limited to, the mouse

-21-

fibroblast cell NIH 3T3, CHO cells, Ltk⁻ cells, etc. Expression plasmids such as that described supra may be used to transfect mammalian cells by methods well known in the art such as calcium phosphate precipitation, electroporation.

This invention also provides a method for producing a polypeptide (e.g. acylcoenzyme A: cholesterol acyltransferase) which comprises growing a host vector system under suitable conditions permitting production of the polypeptide and recovering the polypeptide so produced. Methods of recovering polypeptides produced in such host vector systems are well-known in the art and typically include steps involving cell lysis, solubilization and chromatography.

This invention also provides a method of obtaining a polypeptide in purified form which comprises: (a) introducing a vector, as described above, into a suitable host cell; (b) culturing the resulting cell so as to produce the polypeptide; (c) recovering the polypeptide produced in step (b); and (d) purifying the polypeptide so recovered. As discussed above the vector may include a plasmid, cosmid, yeast artificial chromosome, bacteriophage or eukaryotic viral DNA. Also, the host cell may be a bacterial cell (including gram positive cells), yeast cell, fungal cell, insect cell or animal cell. Suitable animal cells include, but are not limited to HeLa cells, Cos Cells, CV1 cells and various primary mammalian cells. Culturing methods useful for permitting transformed or transfected host cells to produce polypeptides are well known in the art as are the methods for recovering polypeptides from such cells and for purifying them.

35

-22-

Using the aforementioned method, this invention also provides a purified wildtype acylcoenzyme A: cholesterol acyltransferase II or III and a purified mutant acylcoenzyme A: cholesterol acyltransferase II or III.

5

This invention also provides an oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or III without hybridizing to a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II or III. Further, this invention also provides an oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within the nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II or III without hybridizing to a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or III. These oligonucleotide DNA or RNA. Such oligonucleotides may be used in accordance with well known standard methods for known purposes, for example, to detect the presence in a sample of DNA which will hybridize thereto.

25 The oligonucleotides include, but are not limited to, oligonucleotides that hybridize to mRNA encoding acylcoenzyme A: cholesterol acyltransferase II or III so as to prevent translation of the protein.

30 This invention also provides a nucleic acid having a sequence complementary to the sequence of the isolated nucleic acid which encodes acylcoenzyme A: cholesterol acyltransferase II or III.

35 This invention also provides a method for determining

-23-

whether a subject known to have an imbalance in sterol levels has the imbalance due to a defect in esterification of sterol which comprises (a) obtaining from the subject an appropriate sample containing a mixture of all of the subject's nucleic acids; and (b) determining whether any nucleic acid in the sample from step (a) is, or is derived from, a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase so as to thereby determine whether the subject's imbalance in sterol levels is due to a defect in esterification of sterol. The determination step (b) may comprises: (i) contacting the sample of step (a) with the isolated nucleic acid which encodes acylcoenzyme A: cholesterol acyltransferase II or III or the oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or III without hybridizing to a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II or III under conditions permitting binding of any nucleic acid in the sample which is, or is derived from, a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase to the nucleic acid or oligonucleotide so as to form a complex; (ii) isolating the complex so formed; and (iii) identifying the nucleic acid in the isolated complex so as to thereby determine whether any nucleic acid in the sample contains a nucleic acid which is, or is derived from, a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II or III. In this method, both the isolation of any complex formed are effected using standard methods well known in the art.

In order to facilitate identification of the nucleic acid

-24-

from step (a) the isolated nucleic acid or the oligonucleotide is labeled with a detectable marker. The detectable marker may be a radioactive isotope, a fluorophore or an enzyme. In additions, the nucleic acid sample may be bound to a solid matrix before performing step (I).

This invention also provides a method for treating a subject who has an imbalance in sterol levels due to a defect in esterification of sterol which comprises introducing an isolated nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or III into the subject under conditions such that the nucleic acid expresses a wildtype acylcoenzyme A: cholesterol acyltransferase II or III, so as to thereby treat the subject.

This invention also provides a method for inhibiting wildtype acylcoenzyme A: cholesterol acyltransferase II or III in a subject which comprises transforming appropriate cells from the subject with a vector which expresses the nucleic acid complementary to the isolated nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or III, and introducing the transformed cells into the subject so as to thereby inhibit wildtype acylcoenzyme A: cholesterol acyltransferase II or III. Further, in a preferred embodiment, the nucleic acid is capable of specifically hybridizing to a mRNA molecule encoding acylcoenzyme A: cholesterol acyltransferase II or III so as to prevent translation of the mRNA molecule.

This invention also provides a method for inhibiting the wildtype acylcoenzyme A: cholesterol acyltransferase II or III in a subject which comprises introducing an

-25-

oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II or III without hybridizing to a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II or III into the subject so as to thereby inhibit the wildtype acylcoenzyme A: cholesterol acyltransferase II or III. The oligonucleotide is capable of specifically hybridizing to a mRNA molecule encoding acylcoenzyme A: cholesterol acyltransferase II or III so as to prevent translation of the mRNA molecule.

This invention also provides for a method for identifying a chemical compound which is capable of inhibiting acylcoenzyme A: cholesterol acyltransferase II or III in a subject which comprises (a) contacting a wildtype acylcoenzyme A: cholesterol acyltransferase II or III with the chemical compound under conditions permitting binding between the acylcoenzyme and the chemical compound (b) detecting specific binding of the chemical compound to the acylcoenzyme; and (c) determining whether the chemical compound inhibits the activity of the coenzyme so as to identify a chemical compound which is capable of inhibiting acylcoenzyme A: cholesterol acyltransferase II or III in a subject.

This invention also provides method for differentially inhibiting one acylcoenzyme A: cholesterol acyltransferase but not others using the above methods. In an embodiment, only acylcoenzyme A: cholesterol acyltransferase I is inhibited. In another embodiment only acylcoenzyme A: cholesterol acyltransferase II (ARGP1) is inhibited. In an another embodiment only acylcoenzyme A: cholesterol acyltransferase III (ARGP2)

-26-

is inhibited. Alternatively, two of the acylcoenzyme A: cholesterol acyltransferases may be inhibited. This invention further provides pharmaceutical compositions which will differentially inhibit one or more
5 acylcoenzyme A: cholesterol acyltransferases.

This invention also provides for a pharmaceutical composition comprising the chemical compound identified by the above-described method in an amount effective to
10 inhibit acylcoenzyme A: cholesterol acyltransferase II or III in a subject and a pharmaceutically effective carrier.

This invention also provides a method of treating a
15 subject who has atherosclerosis comprising the above-described pharmaceutical composition. A method of treating a subject who has hyperlipidemia comprising the above-described pharmaceutical composition.

This invention also provides a transgenic, nonhuman mammal comprising the isolated nucleic acid which encodes acylcoenzyme A: cholesterol acyltransferase II or III. The mammal includes, but is not limited to, a mouse, bovine, cat or dog.
25

This invention is illustrated in the Experimental Details section which follows. These sections are set forth to aid in an understanding of the invention but are not intended to, and should not be construed to, limit in any
30 way the invention as set forth in the claims which follow thereafter.

-27-

Experimental Details

First Series of Experiments

Example 1:5 Materials and Methods:

Transformation of yeast was performed with lithium acetate (15) by amino-acid prototrophy selection. A diploid strain (5051) was constructed between two isogenic derivatives of W303 (16); W1346-3C (MATa, ade2-1, can1-100, his3-11, 15, leu2-3, 112, trp1-1, ura3-1) and W1134-2C (MATa, can1-100, his3-11, 15, leu2-3, 112, trp1-1, ura3-1, met14DHpaI-SalI). Growth on complete (YEPD) or synthetic medium, sporulation and dissection was performed as described (17).

Competent cells of *Escherichia coli* strain DH5a (Gibco-BRL) and DNA modifying enzymes (Promega) were used according to the manufacturers instructions. pH3(34), from L.A. Grivell, was digested with Nhe I, blunt-ended with Klenow sequences, and digested with Avr II to liberate a 1614-bp fragment. An Xba I, Sma I fragment of pJH-H1 encoding the *HIS3* gene was then inserted at these sites in the vector backbone to produce the *are1ΔNA* allele. This construct was digested with Bsa I to liberate a 3821-bp fragment which was then transformed into strain 5051. Disruption of *ARE1* was confirmed by Southern blot analysis.

30 Radioactive probes of acylcoenzyme A: cholesterol acyltransferase-related enzyme I were prepared by random priming (Pharmacia) with ³²P-dCTP. Genomic DNA (18) was transferred to Hybond membranes (Amersham) and hybridized in the absence of formamide at 65° or 60°C (19).

-28-

A shotgun library of cosmid 14-21 from chromosome XIV (Peter Philippsen, Biozentrum Basel) was constructed using the nebulizing technique (20). The DNA was nebulized (90 seconds, 2 bars), size fractionated, treated with DNA polymerase I (Klenow fragment) and T4 DNA polymerase and blunt-end ligated into pTZ18R (Pharmazia, Germany). Nucleotide sequencing was performed by dideoxy-chain-termination with digoxigenin-labeled reverse primer and Sequenase (United States Biochemical). The reactions were analyzed on the GATC 1500 direct blotting electrophoresis system (GATC GmbH, Germany) using the Boehringer-Mannheim Dig-development protocol. Sequences were aligned by SeqMan (DNA Star Inc.). Database searching was performed with BLAST (21) and GCG Inc. software (22). The DNA sequence of the acylcoenzyme A: cholesterol acyltransferase-related enzyme I and acylcoenzyme A: cholesterol acyltransferase-related enzyme II genes are deposited at GenBank (P25628 and U51790, respectively).

20

K O - 5 a n d K O - 3 ' p r i m e r s
(GAGGGGACGAAAATTAGCCGCTATTAATTCTGGTATTGCCACCTAGACAAGAAG
TAAACAGACACAGATGcaagagttcgaatctcttagc (Sequence ID No.:
15) and CTATAAAGATTTAATAGCTCCACAGAACAGTTGCAGGATGCCTTAGGGT
CGActacgtcgttaaggccggtttctgac (Sequence ID No.: 16),
respectively; lower case corresponds to the *LEU2* gene)
were used in a PCR with the *LEU2* gene as a template to
produce the selectable yeast gene flanked by acylcoenzyme
A: cholesterol acyltransferase-related enzyme II gene
sequences (23). This was used to transform a derivative
of yeast strain 5051, heterozygous for the *are1ΔNA*
allele. To identify integrants at the acylcoenzyme A:
cholesterol acyltransferase-related enzyme II locus, a
PCR was performed on genomic DNA from these strains using

-29-

are2-5' (CATTGCAGTTACACGTGAATGC) (Sequence ID No.: 17),
are2-3': (TAGCTCCACAGAACAGTTGCAGG) (Sequence ID No.: 18)
and a 3' ampimer corresponding to the *LEU2* gene (L2-3'
CTCTGACAACAACGAAGTCAG) (Sequence ID No.: 19).

5

1-2 units at an absorbance of 600nm of cells were
incubated in YEPD or defined media containing 1 μ Ci/ml
³H-oleate in tyloxapol/ethanol (1:1) for 16 hours. Total
10 lipids were prepared by hexane extraction (25) and
analysed by thin layer chromatography on DC-plastikfolien
kieselgel 60 plates (E-Merck, Germany). The plate was
developed in hexane, diethyl ether and acetic acid
(70:30:1) and stained with iodine vapor. Incorporation
15 of label into triglyceride and ergosterol ester was
ascertained following scintillation counting and
normalization to a ¹⁴C-cholesterol internal standard and
the dry weight of the cells.

20 To overexpress the acylcoenzyme A: cholesterol
acyltransferase-related enzyme I gene by copy number
under the control of its own promoter in YEp3-16, a 2354
bp Cla I fragment from pH3(34), encompassing the entire
acylcoenzyme A: cholesterol acyltransferase-related
25 enzyme I gene, was made blunt-ended with Klenow DNA
polymerase I and introduced into the Sma I site of
YEp352. To constitutively overexpress acylcoenzyme A:
cholesterol acyltransferase-related enzyme I from the ADH
promoter in pADH5-36, a 2290 bp Nar I fragment of
30 pH3(34), starting 70 bp 5' to the ORF was blunt-ended
with Klenow and ligated to Klenow-treated, Eco RI
digested, pDC-ADH (a derivative of pS5) (26). Increased
expression of the acylcoenzyme A: cholesterol
acyltransferase-related enzyme I transcripts, relative to
35 a wild-type cell, was confirmed by northern blot

-30-

analysis.

The incorporation of [1-¹⁴C] acetate into saponified lipids was assessed as a measurement of sterol synthesis. Approximately 2 OD₆₀₀ units of cells were incubated with 20 μ Ci [1-¹⁴C] acetate in 2 ml defined media at 30°C for 3 hours and subjected to lipid saponification, hexane extraction and TLC chromatography (29). The incorporation of counts into total sterols were assessed following scintillation counting. To normalize the estimate of sterol biosynthesis to incorporation of acetate into the fatty acid pool, the aqueous lysate remaining after hexane extraction was acidified with concentrated HCl and re-extracted with hexane (30).

Experimental Discussion

To use yeast genetics to study sterol esterification, the human acylcoenzyme A: cholesterol acyltransferase sequence was used to search for homologous yeast genes and subsequently to identify an additional human isoform (Figures 1A and 1B). Acylcoenzyme A: cholesterol acyltransferase related enzyme I, an 1830-bp open reading frame (ORF) on yeast chromosome III, encodes a 610-residue protein with 23% identity and 49% similarity to human acylcoenzyme A: cholesterol acyltransferase I (Figures 1A and 1B). The yeast and human proteins possess leucine zipper motifs that could mediate protein-protein interactions (esterification is probably performed by a multimeric complex) (12), and possess at least two predicted transmembrane domains that may mediate the membrane association of the acylcoenzyme A: cholesterol acyltransferase reaction (13, 14).

To define the role of acylcoenzyme A: cholesterol

-31-

acyltransferase-related enzyme I in sterol esterification, the deletion mutant, *are1 Δ NA*, was generated by homologous recombination (15, 16, 17) (Fig. 2A). In a diploid strain, a 1614-bp segment of one acylcoenzyme A: cholesterol acyltransferase-related enzyme I allele was replaced with the *HIS3* gene and confirmed by Southern hybridization (Fig. 2B). Analysis of mutant and wild-type haploid progeny from this diploid indicated no differences in growth rates or incorporation of ³H-oleate into ergosterol ester.

The lack of a defect in sterol esterification in *are1 Δ NA* strains could result from alternate esterification activities. Reduced stringency hybridization of yeast genomic DNA with the acylcoenzyme A: cholesterol acyltransferase-related enzyme I coding sequence as a probe indicated that additional homologous sequences were present (18, 19). A Bam HI digestion of genomic DNA produced the predicted 2.9-kb acylcoenzyme A: cholesterol acyltransferase-related enzyme I fragment and a ~6.0-kb hybridizing fragment (Fig. 2C). Contour clamped homogeneous electric field electrophoretic analysis of yeast chromosomes suggested the latter sequence was localized to chromosome X or XIV. On the basis of homology to acylcoenzyme A: cholesterol acyltransferase-related enzyme I, this gene, designated acylcoenzyme A: cholesterol acyltransferase-related enzyme II, encodes a second yeast homolog to human acylcoenzyme A: cholesterol acyltransferase I (Figures 1A and 1B). The genomic sequence (20, 21, 22) encompassing acylcoenzyme A: cholesterol acyltransferase-related enzyme II on chromosome XIV predicts a 5997-bp Bam HI fragment and a 1929-bp ORF, which translates into a 643-residue polypeptide. The yeast acylcoenzyme A: cholesterol

-32-

acyltransferase related enzymes genes are 61% and 49% identical at the DNA and predicted protein levels, respectively. Arelp, Are2p and the human acylcoenzyme A: cholesterol acyltransferaseI protein are most related at the COOH-terminal region (42% identity over a 90-residue sequence) (Figures 1A and 1B).

To assess the contribution of Are2p to sterol esterification, one copy of the acylcoenzyme A: cholesterol acyltransferase-related enzyme II coding sequence was deleted from the genome of an *ARE1/are1ΔNA* heterozygous diploid by a polymerase chain reaction approach (23) (Fig. 2D). Haploid progeny representing the single *are1ΔNA* and *are2Δ* deletions and the *are1ΔNAare2Δ* double mutant were obtained. To ascertain the effect of deletion of acylcoenzyme A: cholesterol acyltransferase-related enzymes genes upon cytoplasmic lipid storage, the neutral lipid components (triglyceride and sterol ester) of the yeast cells were detected by fluorescence microscopy after staining with Nile Red (24). In wild-type cells, cytoplasmic fluorescent droplets accumulated in stationary phase cultures (Fig. 3A). No differences in are single mutants were detected. However, the number of droplets observed in *are1ΔNAare2Δ* double mutants, was one-third to that in wild-type strains (Fig. 3B; over multiple fields, 5.57 ± 2.73 vs. 16.73 ± 4.6 droplets/cell, $P < 0.05$).

The wild-type and are mutant cells were analyzed for the incorporation of ^3H -oleate into sterol ester (25) (Fig. 4B). No significant differences in triglyceride biosynthesis were detected. In contrast to normal sterol ester biosynthesis observed in *are1ΔNA* mutants, deficiencies in sterol esterification were apparent in

-33-

both *are2Δ* and *are1ΔNAare2Δ* mutants. These were detected by iodine vapor staining of thin layer chromatographs of total yeast lipids in addition to the oleate incorporation assays. Sterol ester levels of *are2Δ* single mutants were reduced to less than 26% of wild-type suggesting the acylcoenzyme A: cholesterol acyltransferase-related enzyme II isoform to confer the majority of acyltransferase activity. The *are1ΔNAare2Δ* double mutant was almost totally deficient in sterol esterification (less than 1% of wild-type levels). In confirmation of the critical role of Are proteins in sterol esterification, microsomes from double mutant yeast cells lacked acylcoenzyme A: cholesterol acyltransferase activity when assayed *in vitro*.

To confirm that the protein encoded by an acylcoenzyme A: cholesterol acyltransferase-related enzymes ORF was sufficient for sterol esterification, the acylcoenzyme A: cholesterol acyltransferase-related enzyme I coding sequence was over-expressed in vectors with increased copy number (YE_p3-16) or elevated transcription (the alcohol dehydrogenase promoter in pADH5-36) (26). There were no detectable changes in triglyceride or phospholipid biosynthesis resulting from acylcoenzyme A: cholesterol acyltransferase-related enzyme I over-expression. In *are2Δ* or *are1ΔNAare2Δ* double mutants, acylcoenzyme A: cholesterol acyltransferase-related enzyme I over-expression complemented the sterol esterification defect (Fig. 4C). In wild-type and *are1ΔNA* single mutants, the high level expression of acylcoenzyme A: cholesterol acyltransferase-related enzyme I did not elevate sterol ester synthesis above untransformed controls. This suggests that either substrates are limiting in acylcoenzyme A: cholesterol

-34-

acyltransferase-related enzymes strains or that the enzyme is post-translationally regulated as in mammalian cells (27).

5 An accumulation of unesterified sterol in cell membranes would likely be deleterious (28). However, despite the major changes in sterol esterification conferred by the are mutants, we did not detect any reduction in growth rates. The established role of sterol esterification in
10 the storage of sterol suggests that an inability to esterify sterol could lead to homeostatic changes in sterol biosynthesis. This relationship might account for the viability of the mutants. Total lipids, labelled by the incorporation of [1-¹⁴C] acetate into exponentially
15 growing cells (29, 30), were saponified and extracted. The *are1 ΔNaare2Δ* double mutants had a two to three-fold lower level of sterol biosynthesis than wild-type cells, although no changes were observed in the single mutants (Fig. 4D). In fact, free sterol concentrations were
20 roughly equivalent in all cells. Feedback regulation of sterol biosynthesis by acylcoenzyme A: cholesterol acyltransferase activity has been observed in mammalian cells (31) and may be a common mechanism that maintains intracellular sterol at non-toxic concentrations.

25 The involvement of multiple gene families in sterol homeostasis is common in mammalian and yeast cells, for example, the LDL receptor related protein and scavenger receptor gene families, the SREBP family, and 3-hydroxy-
30 3-methyl-glutanyl-CoA reductase) (4, 32, 33, 34). This apparent redundancy of function has clear physiological consequences as evidenced by deletion of any one of the family members. The observation here of two yeast genes for sterol esterification provoked the hypothesis of

-35-

similar redundancy for this reaction in humans. To this end, a consensus of the yeast acylcoenzyme A: cholesterol acyltransferase-related enzymes and human acylcoenzyme A: cholesterol acyltransferaseI sequences was used to
5 identify an additional cDNA with significant identity (47%) to human acylcoenzyme A: cholesterol acyltransferaseI and the yeast proteins (Figure 1B, Genbank accession # R07932).

10 Sterol homeostasis is a complex event under subtle regulatory controls, one component of which is sterol esterification. The demonstration here of multiple yeast and human acylcoenzyme A: cholesterol acyltransferase isoforms raises the possibility that *in vivo*, the enzymes
15 exhibit alternate substrate preferences. The analysis of esterification reactions in yeast is likely to impact the understanding of sterol homeostasis and atherosclerosis in humans.

20 Example 2:

Tissue specific expression of acylcoenzyme A: cholesterol acyltransferase II was analyzed by Northern blot RNA hybridization of RNA obtained from the described tissues.
25 Using the same materials and procedures of Chang, et al. (10), the specific expression of acylcoenzyme A: cholesterol acyltransferase II in liver and muscle is documents, in contrast to similar experiments using the previously known acylcoenzyme A: cholesterol
30 acyltransferase I (10) (Figures 7A and 7B). Acylcoenzyme A: cholesterol acyltransferase II was also detected and specifically expressed in adrenal, thyroid and testicular tissues.

35

-36-

Example 3:

After determining the consensus sequence between the two yeast gene and the previously known human acylcoenzyme A: cholesterol acyltransferase, the consensus sequence was compared to sequences deposited in Genbank. The clones containing the sequences that showed similarity to the consensus sequence were ordered from the I.M.A.G.E. Consortium, affiliated with Research Genetics, Inc., 2130 Memorial Parkway S.W. Huntsville, Alabama 35801. Clones deposited with the I.M.A.G.E. consortium are publicly available upon request. A particular clone, Genbank ID clone No. Z39933 was chosen. This clone contains a cDNA fragment whose sequence encodes human acylcoenzyme A: cholesterol acyltransferase II. The fragment was cut out with restriction enzymes Bgl II and Not I. The resulting fragment was introduced into the yeast expression vector pRS426 at Bgl II and Not I sites downstream of the yeast promoter (GAL1/GAL10) which is regulated by carbon sources. The resultant vector was designated YepAB-ACAT2 (Figure 6).

Example 4:

Antisense RNA technology can be used to create mice, or mouse or human cell lines incapable of translating acylcoenzyme A: cholesterol acyltransferase II RNA into protein. Standard methods may be used to create an antisense oligonucleotide to the human homolog of acylcoenzyme A: cholesterol acyltransferase II. These methods are well known in the art (36).

Specifically, part or all of a wildtype acylcoenzyme A: cholesterol acyltransferase II is ligated adjacent to a mammalian promoter in the opposite orientation. The

-37-

promoter and other replicatory mechanisms inside the cell will transcribe a human homolog of acylcoenzyme A: cholesterol acyltransferase II encoding, nonsense strand. This strand will bind with the coding mRNA which is normally synthesized to form a complex. Due to the formation of this complex, the antisense strand prevents the translation of the coding mRNA into protein.

Further, one skilled in the art can synthesize an oligonucleotide *in vitro* which is capable of binding the mRNA that encodes a human homolog of acylcoenzyme A: cholesterol acyltransferase II so as to inhibit the translation of the mRNA into protein. The oligonucleotides can then be introduced into the subject using a pharmaceutically acceptable carrier. Methods of synthesizing naturally and non-naturally occurring oligonucleotides which are capable of inhibiting the translation of the mRNA into protein are well known in the art. Also, means of transfecting an organism with such oligonucleotides are well known in the field.

Example 5:

Mice can be made with an alteration in their genome, specifically at the acylcoenzyme A: cholesterol acyltransferase II gene site. Standard methods may be used to alter the genome. These methods are well known in the art (37, 38).

One such process to achieve this goal involves disrupting the wildtype mouse homolog of acylcoenzyme A: cholesterol acyltransferase II *in vitro*, then introducing the altered gene into mouse embryonal stem cells in such a way as to target integration into the corresponding genomic region. This process can be performed such that both copies of

-38-

- the wildtype acylcoenzyme A: cholesterol acyltransferase II are replaced by the altered, knock-out version. These modified cells can be introduced into blastocysts which will be allowed to develop into chimeric adults. Mice
- 5 bearing the altered acylcoenzyme A: cholesterol acyltransferase II gene will be mated to each other to generate homozygous mutant acylcoenzyme A: cholesterol acyltransferase II animals.
- 10 Further, one can breed two mice who are heterozygous for mutant acylcoenzyme A: cholesterol acyltransferase II. From their progeny, one skilled in the art could select the progeny who are homozygous for mutant acylcoenzyme A: cholesterol acyltransferase II. Breeding and selecting
- 15 such progeny are well known in the art.

References of the First Series of Experiments

1. T.-Y. Chang and G.M. Doolittle, in *The Enzymes*, P. Boyer, Eds. (Academic Press, New York, 1983), vol. XVI, pg. 523.
- 5 2. K. E. Bloch, *CRC Crit. Rev. Biochem.* **14**, 47 (1983).
3. M. P. Reinhart, J. T. Billheimer, J. R. Faust, J. L. Gaylor, *J. Biol. Chem.* **262**, 9649 (1987).
4. X. Wang et al., *Cell* **77**, 53 (1994).
5. J. T. Billheimer and P. J. Gillies, in *Advances in Cholesterol Research* M. Esfahani, J. B. Swaney, Eds. (Telford Press, Philadelphia, 1992).
- 10 6. T. P. Carr, J. S. Parks, L. L. Rudel, *Arterio. and Thromb.* **12**, 1275 (1992).
7. S. C. Rumsey, N. F. Galeano, B. Lipschitz, R. J. Deckelbaum, *J. Biol. Chem.* **270**, 10008 (1995).
- 15 8. T. Shiratori and D. S. Goodman, *Biochem. Biophys. Acta* **106**, 625 (1965).
9. S. S. Katz, G. G. Shipley, D. M. Small, *J. Clin. Inv.* **58**, 200 (1976).
- 20 10. C. C. Y. Chang, H. Y. Huh, K. M. Cadigan, T. Y. Chang, *J. Biol. Chem.* **268**, 20747 (1993).
11. D. Cheng, C. C. Y. Chang, X. M. Qu, T. Y. Chang, *J. Biol. Chem.* **270**, 685 (1994).
12. S. K. Erickson, S. R. Lear, M. J. McCreery, *J. Lipid Res.* **35**, 763 (1994).
- 25 13. D. S. Goodman, D. Deykin and T. Shiratori, *J. Biol. Chem.* **239**, 1335 (1964).
14. G. M. Doolittle and T. Y. Chang, *Biochem.* **21**, 674 (1982).
- 30 15. H. Ito, Y. Fukuda, K. Murata, A. Kimura, *J. Bacteriol.* **153**, 163 (1983).
16. B. J. Thomas and R. Rothstein, *Cell* **56**, 619 (1989).
17. F. Sherman and J. Hicks, in *Methods in Enzymology* C. Guthrie, G. R. Fink, Eds. (Academic Press, 1991),

- vol. 194, pp. 21.
18. C. S. Hoffman and F. Winston, *Gene* **57**, 267 (1987).
 19. F. M. Ausubel et al., *Current Protocols in Molecular Biology* (John Wiley & Sons, New York, 1987), vol. 1.
 - 5 20. C. M. Okpodu et al., *BioTechniques* **16**, 154 (1994).
 21. S. F. Altschul et al., *J. Mol. Biol.* **215**, 403 (1990).
 22. J. Devereux, P. Haeberli, O. Smithies, *Nucl Acids Res* **12**, 387 (1984)
 - 10 23. A. Baudin, O. Ozier-Kalogeropoulos, A. Denouel, C. Cullin, *Nuc. Acids Res.* **21**, 3329 (1993).
 24. P. Greenspan, E. P. Mayer, S. D. Fowler, *J. Cell Biol.* **100**, 965 (1985).
 25. L. W. Parks, C. D. Bottema, R. J. Rodriguez, T. A. Lewis, *Methods Enzymol* **111**, 333 (1985).
 - 15 26. S. L. Sturley et al., *J. Biol. Chem.* **269**, 21670 (1994).
 27. C. C. Chang et al., *J. Biol. Chem.* **270**, 29532 (1995).
 - 20 28. G. J. Warner et al., *J. Biol. Chem.* **270**, 5772 (1995).
 29. R. Y. Hampton and J. Rine, *J. Cell Biol.* **125**, 299 (1994).
 30. D. Dimster-Denk, M. K. Thorsness, J. Rhine, *Mol. Biol. Cell* **5**, 655 (1994).
 - 25 31. I. Tabas, D. A. Weiland, A. R. Tall, *J. Biol. Chem.* **261**, 3147 (1986).
 32. M. Krieger and J. Herz, *Annu. Rev. Biochem* **63**, 601 (1994).
 - 30 33. X. Wang et al., *Cell* **77**, 53 (1994).
 34. M. E. Basson, M. Thorsness, J. Rine, *Proc. Natl. Acad. Sci. USA* **83**, 5563 (1986).

-41-

35. S. L. Thompson, R. Burrows, R. J. Laub, S. K. Krisans, *J. Biol. Chem.* **262**, 17420 (1987).
36. Griffin, E.F. and H. Harris (1992) *J. Cell Science* **102**: 799-805.
- 5 37. Joyner, A.L. and F. Guillemot (1994) *Current Opinion in Neurobiology* **4**: 37-42.
38. Duyao, M.P., et al. (1995) *Science* **269**: 407-410.

10

-42-

5 cond Series of Experiments

The efficient regulation of intracellular sterol levels is required for cell viability by all eukaryotic organisms. When this regulation is aberrant in cells of the arterial wall, disease states such as atherosclerosis ensue. A critical component of this homeostasis is intracellular sterol esterification reaction, mediated by the enzyme, acyl coenzyme A-cholesterol acyltransferase (ACAT). In the model eukaryote, yeast, this laboratory has demonstrated that sterol esterification is mediated by a two gene family (Yang et al., Science, 1996, 272:1353). The existence in human cells of two additional genes encoding ACAT related enzymes are demonstrated. These protein are termed ACAT related gene products (ARGP) 1 and 2, also known as acylcoenzyme A: cholesterol acyltransferase II and acylcoenzyme A: cholesterol acyltransferase III respectively. The ARGPs exhibit marked sequence conservation to the human ACAT sequence (hACAT) originally identified by Chang and colleagues. ARGP1 is expressed at high levels in intestine and liver in contrast to the expression of hACAT which is of low abundance in these tissues. The observation that knock-out mutant mice deficient in the murine homolog if hACAT retain sterol esterification activity in liver and intestine (Meiner et al., PNAS, 1996, 93:14041), suggests that ARGP1 is a candidate for sterol esterification in these tissues. The expression of ARGP2, by contrast, seems to be restricted to the fetal liver, suggesting it to have a role in lipid metabolism during development. Analysis of genome databases indicates that ACAT-like gene families are a common occurrence in multiple organisms. It is hypothesize that multiple enzymes for sterol esterification will provide flexibility in response to differing sterol and fatty acid substrates encountered by different tissues. This further suggests specific roles

-43-

for these enzymes in lipoprotein production, lipid homeostasis, and disease progression.

5 The regulation of membrane sterol levels is required for cell viability by all eukaryotic organisms. When this regulation is aberrant in human cells, disease states such as atherosclerosis (excessive accumulation of cellular esterified cholesterol in cells of the arterial wall, reviewed in (1-4)), Niemann Pick C (inability to store sterol correctly, resulting in lysosomal lipidosis, (5)) or Wollmann's disease (a defect in sterol ester hydrolysis, (6)) ensue. A critical component of this homeostasis is the intracellular neutralization of sterol by an esterification reaction between the C₃-OH group of cholesterol and fatty acyl-coenzyme A. This reaction is performed in mammalian cells by the enzyme acyl coenzyme A-cholesterol acyltransferase (ACAT). Since the process of sterol esterification converts sterol into a cytoplasmic storage form, it is critical to all eukaryote, including the microorganism Saccharomyces cerevisiae (budding yeast). Analysis of sterol homeostasis in this model organism has the advantage that molecular genetics, particularly since the completion of the yeast genome sequencing project, is powerful and relatively straightforward. Taking advantage of this, it is demonstrated that sterol esterification in yeast is mediated by a two gene family (7), neither of which is essential for life. These genes (ARE1 and ARE2; encoding ACAT Related Enzymes 1 and 2, respectively) are both capable of independently esterifying sterol, although in terms of contribution to the sterol ester mass of the cell, Are1 is a minor isoform relative to Are2. The genes are structurally and functionally analogous to the ACAT sequence isolated originally from macrophages by Chang and colleagues (8). They share approximately 23% identity at the protein level and expression of the human macrophage ACAT cDNA in yeast are

-44-

double deletion mutants results in esterification of sterol (9).

5 A critical test of the role of the ACAT gene product in cholesterol homeostasis and atherosclerosis was initiated by Farese and colleagues, by the production of "knock-outs" at the Acact locus corresponding to the mouse homolog of hACAT (10). The fidelity of the mutation was confirmed by sequencing of cDNA from the disrupted allele and by the failure to detect immunoreactive protein in Acact-/- cell extracts. The animals were healthy and fertile and had residual, but significant, sterol esterification activity in fibroblasts and macrophages. Cholesterol ester levels and ACAT activity in the adrenals were also severely reduced. Conversely, Acact-/- 15 livers contained significant levels of cholesterol ester, and esterification activity was not altered. Furthermore, sterol absorption in the intestine, a process that probably requires esterification, was unaffected by the gene disruption. These observations 20 strongly suggest that as in yeast, there are multiple genes for the ACAT reaction in mammalian cells, probably with tissue specific expression patterns.

25 Interestingly, despite the clear origin of the yeast gene family by gene duplication, the ARE proteins have diverged such that the majority of sequence conservation is in the COOH-terminal domain of the protein. This is presumably the critical region of the molecule, since it 30 is also conserved in the human protein. Using this region as a database probe, R07932 (11) was identified, a partially sequenced cDNA entry in the database of expressed sequence tags (best); R07932 exhibits significant similarity to the ACATs particularly over the COOH-terminal region. Taken together; the "founder" 35 sequence, the observations in yeast of a two gene family for sterol esterification, and the tissue-specific

-45-

expression patterns of enzyme activity in Acact-knock-out mice, suggest that there are multiple genes for this reaction in all eukaryote. It is reported here the isolation and characterization of cDNAs from two human loci that encode ACAT Related Gene Products (ARGP). ARGP1 is represented multiple times in the best, including R07932, and is expressed ubiquitously with the highest levels occurring in the liver, intestine and adrenal gland. By contrast, sequences identical to ARGP2 in the databases are infrequent, consistent with the observation of an essentially embryonic pattern of expression. Analysis of genome databases indicates that gene families that conserve these motifs are a common occurrence in multiple organism.

15

Materials and Methods.

Database searching for ACAT related sequences. A sequence corresponding to the strongest region of protein conservation between the human macrophage ACAT and yeast ARE sequences was used to identify protein sequences predicted to be encoded by entries in the best using the tblastn software (NCBLI). The DNA sequences thus arising were used to detect additional clones in any available database, that demonstrated overlaps of nucleotide sequence identity. Databases searched included; best, the non-redundant GENBANK, and the confidential database held at The Institute of Genome Research (TIGR). Overlaps between these sequences were detected using the sequence alignment programs, "lineup" and "pileup" from GCG Inc (Madison, WI). A consensus sequence was then generated. Escherichia coli clones with the largest inserts corresponding to these sequences (see table 1) were obtained from the I.M.A.G.E. consortium and resequenced from both ends using commercial primers, T3 and T7, or internal primers derived from a consensus. Nucleotide sequencing was performed at the Columbia University Combined Center core facility using an Applied

35

-46-

Biosystems fluorescent sequencing machine.

Table 1: Entries of human ACAT related gene products in the products in the data base of expressed sequence tags.

5	Gene	Clone ID (IMAGE)	GENEBANK ID	Insert size (bp)	Comments
10	ARGP1	200587	R99213	620	
			R99214		
		55218	c-IMF11	1800	chimera
			Z43867		
			Z33993		
15		1881180	H45923	1000	
			H45924		
		78614	M79086	300	
		153836	R48474	800	
			R48475		
20		106260	T35085	800	
	ARGP2	128921	R10272	680	
			R10273		
		213176	N75438	540	
25		245265	H76642	300	

Isolation and sequencing analysis of full length cDNA clones of ARGP1 and ARGP2. Since in no instance were any of the database clones full length for either ARGP1 or ARGP2, additional clones with intact 5'-ends are described. Several strategies were chosen using a consensus nucleotide sequence derived from the sequencing of the best clones designed and synthesized 3-end, gene specific primers and used a PCR based, rapid amplification of cDNA ends (RACE) to derive 5'-RACE reaction products from a human liver/spleen Marathon library (Clontech®). Similar strategy was used to derive PCR products from a human fetal brain library generously

-47-

provided by Bento Soares (Columbia University). In some instances, a nested PCT reaction was performed using internal gene specific primers and library adaptors. Finally primer extension cDNA products were identified from mRNA extracted from human intestine (a kind gift of P. Dawson). Amplification products of the predicted size were confirmed as gene specific, using southern hybridization to sequences predicted to be at the 3'-end of these products. The products were isolated from agarose gels using Geneclean and subcloned into TA variants of pBluescript (Stratagene®) vectors of klenow/kinase treated and blunt end ligated to pGEM2 (Promega®). Positive clones were identified by colony hybridization or by PCR amplifications using an internal ARGP specific primer. Clones with the largest inserts were sequenced to obtain novel sequence and where necessary, this process was reiterated with ARGP 5' specific primers derived from the new sequence.

Tissue specific expression of hACAT and ARGPs. Fragments of the best clones R99213 and R10273 corresponding to ARGP1 and ARGP2, respectively were derived by digestion with EcoRI and NotI, and purified from agarose gels with Geneclean. A 1.6 kbp fragment corresponding to the human ACAT cDNA identified by Chang et al was used as a probe for the expression of this gene. Radiolabelled probes were generated by random priming (Pharmacia®) in the presence of ³²-P dCTP and used to probe Multiple Tissue Northern (MTN, Clontech®) of human samples. Hybridizations were performed, according to the manufacturers instructions, using ExpressHyb rapid hybridization solution for 1 hour at 78°C, followed by washed in 2xSSC at 55°C and 0.1xSSC, 0.5%SDS at 50°C.

Cell culture expression of ARGPs. To facilitate quantitation of mRNA from the ARGP genes, a reverse-transcriptase PCR (RT-PCR) approach was devised to

-48-

analyze expression in a variety of human (HeoG2, THP-1 macrophages) and rodent (J774 macrophages) and simian (CV1 kidney cells). Where possible, primers were designed to be conserved between rodents and humans (as described below, the mouse sequence homolog to ARGP1) has been identified. Alternatively, PCR conditions were optimized to permit moderate mismatches. The ARGP amplification primers were designed to be gene specific (i.e. to regions not conserve within the family) and to produce distinct size products.

Experimental Results and Discussion

The approach that the region of strongest conservation between the yeast ARE proteins and hACAT would be critical to the function of any sterol esterification enzymes was taken. A region of conservation (consensus; LN---E---FGDR-FY GDWWN, single letter amino-acid code) that is invariant over the three proteins was chosen and a series of entries derived from gene sequencing projects identified. In addition to sequences from Caenorhabditis. elegans, Schizosacharomuces pombe, Drosophila melanogater and Arabidopsis thaliens, several entries in the best of human cDNAs that suggested an independent gene encoding an ACAT like protein were observed. Using the nucleotide sequence to this clone, a second homologous but distinct entry was identified. These proteins are termed, ACAT Related Gene Products (ARGP) 1 (acylcoenzyme A: cholesterol acyltransferase II) and 2 (acylcoenzyme A: cholesterol acyltransferase III). The sequence identified by Chang et al (8) will be referred to as hACAT, hereon. A limited protein sequence to a founder clone (R07932) to ARGP1 has been presented previously (11). The entries in the best that define these two genes, including their insert sizes are described in table 1. As is evident, the majority of inserts (with the exception of a chimeric clone ZA3867) are less than 1Kbp. The northern and sequence analysis

-49-

presented indicated them to be incomplete clones. However, they clearly define two distinct genes of strong similarity to the ACAT sequence, with the majority of predicted protein conservation at the COOH-terminal region. As described below certain motifs considered critical to sterol esterification are conserved. To identify the role of these genes in the reaction, full length ARGP clones were sought and their patterns identified.

ARGP1, a ubiquitously expressed member of the ACAT gene family. To establish the profile of expression of ARGP1, probed multiple tissue northern of human mRNA was probed, using a fragment close to the 3' end of the gene. Although this region displays the maximum conservation at the protein level in this gene family, the genes are sufficiently divergent at the DNA level to be able to design gene specific hybridization probes. The ARGP1 sequence is expressed at abundant levels in may tissues with the exception of lung and kidney. The majority of tissues express a 2.0kb message but, some tissues (e.g. adrenal, small intestine, thymus) also express a 2.4kb mRNA at varying levels. The same northern were hybridized with a probe to the human macrophage ACAT sequence. As described by others(8,12,13), the hACAT sequence detects 4 messages of approximately 3.0, 4.0, 4.7 and 7.4Kb. Upon comparison of the two hybridization results, an overlapping but occasionally differential expression pattern was observed. Adrenal tissues express the highest levels of both hACAT and ARGP1 message. By this analysis, hACAT messages are rare in liver and intestine in contrast to ARGP1 which is highly expressed in these tissues. Conversely, ARGP1 was poorly expressed in kidney, lung and placenta although hACAT mRNA was easily detected. This tissue specific expression suggests that ARGP1 is an ideal candidate for sterol esterification in tissues such as liver and intestine,

-50-

which retain sterol esterification activity in ACAT k/o mice (10).

ARGP2, an embryonic isoform of the ACAT gene family.

5 Efforts to identify a transcript from ARGP2 in adult tissues were unsuccessful. Therefore embryonic tissue samples were chosen to investigate since the original founder clone was derived from a fetal liver library. A multiple tissue northern of mRNA from human embryonic
10 brain, liver, kidney, and lung, were probed with and ARGP2 specific, COOH-terminal probe. As shown in figure 9, a single message of ~2.2kb was identified only in embryonic liver tissues, suggesting a high degree of tissue and developmental specificity to the expression of
15 this gene product.

Expression of ARGP1 in cell culture models. To develop a system in which to test the effect of reaction substrates on the esterification reaction performed by
20 the ARGP enzymes. The expression of these genes in several tissue specific were examined, cell culture models. As shown in figure 10, ARGP1 is clearly expressed in liver (HepG2) and Kidney (CV-1) cell lines. The latter result is somewhat in contrast to the northern
25 blot on human tissue samples. This most likely reflects the sensitivity of the RT-PCR approach compared to filter hybridization and suggests that ARGP1 is probably expressed in most tissues. Alternatively it may represent species difference (simian vs. human) or more
30 interestingly the differentiation status of the cells under study. In data not shown here, ARGP1 was also clearly expressed in human and mouse macrophage models (THP-1 and J774 cells).

35 Sequence characteristics of ARGP1 and ARGP2. By a combination of 5'-RACE and primer extension additional sequence to cDNA s for ARGP1 and ARGP2 (Figs. 11 and 12)

-51-

have been identified. The ARGP1 sequence predicts a 407 amino-acid protein with approximately 27% identity and 52% similarity to the hACAT protein (Fig. 4). Interestingly, as it was observed for the yeast ARE proteins, the strongest conservation exists at the COOH-terminus of the molecules, to the extent that the NH-2-terminal 50% of all these proteins is essentially unrelated sequence. This pattern also persists at the DNA level (not shown). Identification of the genomic sequence to these cDNAs will establish whether this remarkable divergence arises by exon shuffling of common sequences. Alternatively, convergent evolution of domains with conserved functions in sterol esterification or related processes, may have resulted in the generation of these families. Since the level of DNA conservation between ARGP1 and hACAT is quite low (37% identity), the latter possibility seems likely. The conserved regions are discussed in the context of multiple ACAT like sequences below. The ARGP1 sequence predicts a protein of approximately 47kDa with multiple transmembrane domains in similar positions to those predicted in hACAT. This strongly suggests a membrane location for ARGP1 as would be predicted for a sterol esterification enzyme.

ARGP2 displays a significantly higher level of amino acid conservation with hACAT than does ARGP1. Over the sequence shown (Fig. 12), the protein is 59% identical and 79% similar to human ACAT. Over the same region ARGP1 is only conserved at the level of 32% identity. This striking identity is maintained at the DNA level (62% identity) and may suggest that ARGP2 is more closely analogous to hACAT in both its mechanism of action and its origin, than is ARGP1. As for ARGP1, certain hallmark sequences are retained in ARGP2 (see below). The ARGP2 predicted protein also possesses several predicted transmembrane domains. One entry to the best for ARGP2 has also been allocated an STS (sequence tagged

-52-

site) at the Whitehead Institute, (entry # WI-11660) and has thus been mapped to human chromosome 12.

5 Sterol esterification enzymes evolve as gene families in
multiple organisms. Using the hACAT and AGRP nucleotide
sequences as probes of multiple databases, we sought to
establish whether the observation of gene families of
ACAT related enzymes in yeast and humans was a common
10 occurrence in other organisms. In general this is the
case (Fig. 13). Sequences from the genome of C. elegans,
D. melanogaster and S. pombe, have been identified that
are distinct from each other, within an organism, and
exhibit approximately 25% identity at the predicted
protein level. As for all the ACAT-like proteins, the
15 maximum conservation is observed at the COOH-terminal
region, with many of the apparently critical motifs
described below, being maintained. As would be
anticipated the mouse cDNA for ARGP1 exhibits
approximately 85% identity with its human homolog.

20 Sequence conservation between ARGPs and ACAT in multiple
organisms. As described above, these sequences are
ubiquitous. This conservation, across and within
organisms, facilitates the identification of critical
25 domains of esterification enzymes (Fig. 14).
Interestingly, there is no sequence similarity between
any ACAT-like molecule and lecithin cholesterol
acyltransferase (LCAT), despite the shared utilization of
cholesterol. For the hACAT sequence and its murine
30 homologs, a similarity to "signature" motifs of enzymes
involved in acyl adenylation reactions was reported (8,
12). However, these sequences are unlikely to be
critical, since they are not conserved in any homolog
from any other organism. By contrast, there are regions
35 of strong conservation between these molecules which may
be critical to function. In the esterification
defective, SRD4 mutant CHO cell line, the expressed but

-53-

defective ACAT allele encodes a single amino-acid substitution of leucine₂₆₅ lies in a conserved domain of human, rodent and yeast ACAT. Interestingly, this motif in ARGP1 is more degenerate, although the serine is conserved, the flanking sequence is conservatively replaced by similar residues. The ACAT reaction is probably mediated by a multimeric complex, as shown by radiation inactivation experiments (15). Accordingly, the yeast and human sequences all possess "leucine zipper" multimerization motifs. ARGP1 and ARGP2 lack a classical multimerization motif. Although protein phosphorylation as a mode of ACAT regulation has been refuted (16), a very strong region of conservation (consensus over 7 sequences; LN---E---FGDR-FYGDWWN, single letter amino-acid code) predicts a tyrosine kinase consensus motif for phosphorylation. ARGP2 and ARGP1 are no exception to this. In particular the aspartic acid-tryptophan-tryptophan-asparagine (DWWN) sequence appears to be invariant (with the exception of S.pombe, where it is AWWN) and may represent an active site for the esterification reaction. These regions of conservation are targets for mutagenesis and in preliminary experiments appear critical to the activity of the ACAT and ARE enzymes (17).

Why ACAT gene families? The role, if any, of these ACAT sequence homologs in sterol homeostasis is unclear. Since mouse macrophage ACAT is not critical to sterol esterification in the liver and intestine, it is possible that the additional enzymes evolved to recognize alternate substrates and thus promote sterol absorption in the intestine or production of lipoproteins by the liver (39). Future experiments will be directed to complete the molecular characterization of these genes and test these hypotheses.

-54-

References of the Second Series of Experiments

1. Liscum, L., and Underwood, K.W. (1995) J. Biol. Chem. **270**(26), 15443-15446
2. Chang, T.Y., Chang, C.C.Y., and Cadigan, K.M. (1994) Trends Cardiovascul. Med. **4**, 223-230
3. Tabas, I. (1995) Curr. Op. in Lipidol. **6**(5), 260-268
4. Libby, P., and Clinton, S.K. (1993) Curr. Opin. Lipidol. **4**, 3355-363
5. Pentchev, P.G., Brady, R.O., Blanchette-Mackie, E.J., Vanier, M.T., Carstea, E.D., Parker, C.C., Goldin, E., and Roff, C.F. (1994) Biochim. Biophys. Acta **1225**, 235-243
6. Goldstein, J.L., Dana, S.E., Faust, J.R., Beaudet, A.L., and Brown, M.S. (1975) Journal of Biological Chemistry **250**(21), 8487-95
7. Yang, H., Bard, M., Bruner, D.A., Gleeson, A., Deckelbaum, R.J., Aljinovic, G., Pohl, T., Rothstein, R., and Sturley, S.L. (1995) Circulation In Press
8. Chang, C.C.Y., Huh, H.Y., Cadigan, K.M., and Chang, T.Y. (1993) J. Biol. Chem. **268**(28), 20747-20755
9. Yang, H., Cromley, D., Wang, H., Billheimer, J.T., and Sturley, S.L. (1997) J. Biol. Chem. **272**, 3980-3985
10. Meiner, V.M., Cases, S., Myers, H., Snade, E.R., Bellosta, S., Schambelan, M., Pitas, R.E., McGuire, J., Herz, J., and Farese, R.V. (1996) Proc. Natl. Acad. Sci. USA **93**, 14041-14046
11. Yang, H., Bard, M., Bruner, D.A., Gleeson, A., Deckelbaum, R.J., Aljinovic, G., Pohl, T., Rothstein, R., and Sturley, S.L. (1996) Science **272**, 1353-1356
12. Uelman, P.J., Oka, K., Sullivan, M., Chang, C.C.Y., Chang, T.Y., and Chan, L. (1995) J. Biol. Chem. **270**, 26192-26201
13. Pape, M.E., Schultz, P.A., Rea, T.J., DeMattos, R.B., Kieft, K., Bisgaier, C.L., Newton, R.S., and

-55-

- Krause, B.R. (1995) Journal of Lipid Research 36(4), 823-38
14. Cao, G., Goldstein, J.L., and Brown, M.S. (1996) Journal of Biological Chemistry 271(24), 14642-8
- 5 15. Erickson, S.K., Lear, S.R., and McCreery, M.J. (1994) J. Lipid Res. 35, 763-769
16. Corton, J.M., and Hardie, D.G. (1992) European Journal of Biochemistry 204(1), 203-8
- 10 17. Guo, Z., Yang, H., and Sturley, S.L. (1996) Circulation 94, I-35

-56-

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (I) APPLICANT: Stephen L. Sturley
- (ii) TITLE OF INVENTION: DNA ENCODING ACYLCOENZYME A: CHOLESTEROL
ACYLTRANSFERASE 11 AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 19
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE:
(B) STREET: 1185 Avenue of the Americas
© CITY: New York
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10036
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
© OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: Not Yet Known
(B) FILING DATE: Herewith
© CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: John P. White
(B) REGISTRATION NUMBER: 28,678
© REFERENCE/DOCKET NUMBER: 0575/50852-A
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 278-0400
(B) TELEFAX: (212) 391-0525
© TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (I) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3649 base pairs
(B) TYPE: nucleic acid
© STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
GGGTAGAGAC GGGGTTTCAC CGTGTAGCC AGGATGGTCT GGATCTCCTG ACCTCGTGAT      60
CCACCCACCT CGGCCTCCTA AAGTGCTGGG ATTACAGACA TGAGCCACCG CGCCCAGCCC      120
```

-57-

TATTCATCCC	TTTTCAAAAG	TCAGACCCTA	GGAAGCTGGA	GGGAGGTGGG	GCATGGTTTT	180
ACAGTGAATT	TCTGATTTCA	CTCAGGGTGA	TAAATCAGAC	TCTTGGGGAA	GCGGGTGGTG	240
GCTCTGGACA	GCAGCAGGAA	TGGGGATCCA	GTTAGCAACA	AATCCATGGA	CCTATGACAG	300
GCTGAAAGCC	ACCCCTTCTC	CATCTTTGGG	AGGTTGCCAA	TGTCTGATTT	AACACTATCC	360
AATGAATGAT	CATTGAAAGT	AAAAAATAAC	TATCAACTAG	CAGAAAATAT	AAATGGTAAG	420
CATTAGCACA	TATTTACAT	GTTTATATTT	GGCTCTCAGA	TTGACCTATA	AAACAAAGTC	480
TGGGAAATTC	TATATGATCC	TGAAAAAATG	ATACGCTGGT	CTGGATGGTA	GAATAAGTTG	540
GAGAAATGTT	TAAGCCAAAA	TGCAGTCTTA	CCAATGACTT	TTTATTTTAT	TTTATTAATT	600
TTCAGGATTT	TTGGTATACA	GGTGGTTTTT	GGTTACATGG	AAAAGTTCTT	TACTGGTGAT	660
TTCTGAGATT	TTAGTTCACC	CCTTATCCTG	AGCAGTGTAC	ACTGTTCCCA	ATATGTAGCC	720
TTTTATCCCT	CACCCCCTCT	AAGTTCAAGA	AGACTATGGT	CCTGCAGAAA	GCTTTATATG	780
TAATTAACAT	ATCTTTATCT	TTATCTTTAT	AGGCAGTAGA	CTCATCTTTT	GAAACAGATT	840
CCATTAAGAG	TGAATGTGTA	CCCTCCCTCT	AGCCTTTATT	ATTACTGTTT	TTGCTATTAC	900
ATGTGTTAGT	GTATGTGAAT	TTAATGCTTA	AAAATGTATC	CCATTGGCTA	CTATGGCAAA	960
AGGTTGACTC	ATAAGAGTTT	AGCACGGGTT	AAGATCTGAA	AGTTTTCTCC	CAGCCTCTTA	1020
TCACTGGCGC	AGACTTCACA	ATTCATGGAA	GCCACCAGTG	AGATGACATT	GCCTCAGGCA	1080
GTTACTATTT	TTATATTCTA	TAACTCGAGG	AGCTCAGGGT	TTCGGAAATC	ATTAAACTTT	1140
TTTTGTCCTT	TTAAAGTTGG	AGACAGCAAT	TGTAGACAGC	CTTCCAGTGG	GTTATCTTTT	1200
TGTGTCTCCT	TACCTGTGGA	GAAGCCTATT	AGCTGGGATA	TGTAGTTAAA	TAGCTATATT	1260
TATATATATC	CAGGGCACCC	CGAATTGCGG	AGAGCTTCCC	GGAGTCGACC	TTCCTGCTGG	1320
CTGCTCTGTG	ACCGCTTCCC	GGCTCTGCCC	TCTTGGCCGA	AGTGCCCGCT	GCCGGGCGCG	1380
GGCCTCAGAC	AATACAATGG	TGGGTGAAGA	GAAGATGTCT	CTAAGAAACC	GGCTGTCAAA	1440
GTCCAGGGAA	AATCCTGAGG	AAGATGAAGA	CCAGAGAAAC	CCTGCAAAGG	AGTCCCTAGA	1500
GACACCTAGT	AATGGTCGAA	TTGACATAAA	ACAGTTGATA	GCAAAGAAGA	TAAAGTTGAC	1560
AGCAGAGGCA	GAGGAATTGA	AGCCATTTTT	TATGAAGGAA	GTTGGCAGTC	ACTTTGATGA	1620
TTTTGTGACC	AATCTCATTG	AAAAGTCAGC	ATCATTAGAT	AATGGTGGGT	GCGCTCTCAC	1680
AACCTTTTCT	GTTCTTGAAG	GAGAGAAAAA	CAACCATAGA	GCGAAGGATT	TGAGAGCACC	1740
TCCAGAACAA	GGAAAGATTT	TTATTGCAAG	GCGCTCTCTC	TTAGATGAAC	TGCTTGAAGT	1800
GGACCACATC	AGAACAATAT	ATCACATGTT	TATTGCCCTC	CTCATCTCTC	TTATCCTCAG	1860
CACACTTGTA	GTAGATTACA	TTGATGAAGG	AAGGCTGGTG	CTTGAGTTCA	GCCTCCTGTC	1920
TTATGCTTTT	GGCAAATTTT	CTACCGTTGT	TTGGACCTGG	TGGATCATGT	TCCTGTCTAC	1980
ATTTTCAGTT	CCCTATTTTC	TGTTTCAACA	TTGGCGCACT	GGCTATAGCA	AGAGTTCTCA	2040

- 58 -

TCCGCTGATC CGTTCTCTCT TCCATGGCTT TCTTTTCATG ATCTTCCAGA TTGGAGTTCT	2100
AGGTTTTTGA CCAACATATG TTGTGTTAGC ATATACTG CCACCAGCTT CCCGGTTCAT	2160
CATTATATTC GAGCAGATTC GTTTTGTAAT GAAGGCCAC TCATTTGTCA GAGAGAACGT	2220
GCCTCGGGTA CTAAATTCAG CTAAGGAGAA ATCAAGCACT GTTCCAATAC CTACAGTCAA	2280
CCAGTATTTG TACTTCTTAT TTGCTCCTAC CTTATCTAC CGTGACAGCT ATCCCAGGAA	2340
TCCCACTGTA AGATGGGGTT ATGTCGCTAT GAAGTTTGCA CAGGTCTTTG GTTGCTTTTT	2400
CTATGTGTAC TACATCTTTG AAAGGCTTTG TGCCCCCTTG TTTCGGAATA TCAAACAGGA	2460
GCCCTTCAGC GCTCGTGTTT TGGTCCTATG TGGTATTTAA CTCCATCTTG CCAGGTGTGC	2520
TGATTCTCTT CTTACTTTT TTTGCCTTTT TGCCTGCTG GCTCAATGCC TTTGCTGAGA	2580
TGTTACGCTT TGGTGACAGG ATGTTCTATA AGGATTGGTG GAACTCCACG TCATACTCCA	2640
ACTATTATAG AACCTGGAAT GTGGTGGTCC ATGACTGGCT ATATTACTAT GCTTACAAGG	2700
ACTTTCTCTG GTTTTTCTCC AAGAGATTCA AATCTGCTGC CATGTTAGCT GTCTTTGCTG	2760
TATCTGCTGT AGTACACGAA TATGCCTTGG CTGTTTGCTT GAGCTTTTTT TATCCCGTGC	2820
TGTTTCGTGCT CTTCATGTTT TTTGGAATGG CTTTCAACTT CATTGTCAAT GATAGTCGGA	2880
AAAAGCCGAT TTGGAATGTT CTGATGTGGA CTTCTCTTTT CTTGGGCAAT GGAGTCTTAC	2940
TCTGCTTTTA TTCTCAAGAA TGGTATGCAC GTCGGCACTG TCCTCTGAAA AATCCCACAT	3000
TTTTGGATTA TGTCCGGCCA CGTTCCTGGA CTTGTCGTTA CGTGTTTTAG AAGCTTGGAC	3060
TTTGTTTCCT CTTGTCACT GAAGATTGGG TAGCTCCCTG ATTTGGAGCC AGCTGTTTCC	3120
AGTTGTTACT GAAGTTATCT GTGTTATTTG GACCACTCCA GGCTTTACAG ATGACTCACT	3180
CCATTCCTAG GTCACCTGAA GCCAACTGT TGGAAGTTCA CTGGAGTCTT GTACACTTAA	3240
GCAGAGCAGA ACTTTTTTTT TGGGGCTGGG TGGGGGGAGA AGACCGACTA ACAGCTGAAG	3300
TAATGACAGA TTGTTGCTGG GTCATATCAG CTTTATCCCT TGGTAATTAT ATCTGTTTTG	3360
TTTCTTGACT CTGTCCAATC AGAGAATAAA CATCATAGTT TCTTGCCAC TGAATTAGCC	3420
AAAACACTTA GGAAGAAATC ACTTAAATAC CTCTGGCTTA GAAATTTTTT CATGCACACT	3480
GTTGGAATGT ATGCTAATTG AACATGCAAT TGGGGAAGAA AAAATGTAGA ATGATTTTTG	3540
CTATTTCTAG TAGAAAGAAA ATGTCTGTTT TCCAAAGATA ATGTTATACA TCCTATTTTG	3600
TAATTTTTTT GAAAAAGTT CAATGTTTCT TTTTCCTTAGT TTTTACCTT	3660

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 550 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-59-

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Val	Gly	Glu	Glu	Lys	Met	Ser	Leu	Arg	Asn	Arg	Leu	Ser	Lys	Ser	1	5	10	15
Arg	Glu	Asn	Pro	Glu	Glu	Asp	Glu	Asp	Gln	Arg	Asn	Pro	Ala	Lys	Glu	20	25	30	
Ser	Leu	Glu	Thr	Pro	Ser	Asn	Gly	Arg	Ile	Asp	Ile	Lys	Gln	Leu	Ile	35	40	45	
Ala	Lys	Lys	Ile	Lys	Leu	Thr	Ala	Glu	Ala	Glu	Glu	Leu	Lys	Pro	Phe	50	55	60	
Phe	Met	Lys	Glu	Val	Gly	Ser	His	Phe	Asp	Asp	Phe	Val	Thr	Asn	Leu	65	70	75	80
Ile	Glu	Lys	Ser	Ala	Ser	Leu	Asp	Asn	Gly	Gly	Cys	Ala	Leu	Thr	Thr	85	90	95	
Phe	Ser	Val	Leu	Glu	Gly	Glu	Lys	Asn	Asn	His	Arg	Ala	Lys	Asp	Leu	100	105	110	
Arg	Ala	Pro	Pro	Glu	Gln	Gly	Lys	Ile	Phe	Ile	Ala	Arg	Arg	Ser	Leu	115	120	125	
Leu	Asp	Glu	Leu	Leu	Glu	Val	Asp	His	Ile	Arg	Thr	Ile	Tyr	His	Met	130	135	140	
Phe	Ile	Ala	Leu	Leu	Ile	Leu	Phe	Ile	Leu	Ser	Thr	Leu	Val	Val	Asp	145	150	155	160
Tyr	Ile	Asp	Glu	Gly	Arg	Leu	Val	Leu	Glu	Phe	Ser	Leu	Leu	Ser	Tyr	165	170	175	
Ala	Phe	Gly	Lys	Phe	Pro	Thr	Val	Val	Trp	Thr	Trp	Trp	Ile	Met	Phe	180	185	190	
Leu	Ser	Thr	Phe	Ser	Val	Pro	Tyr	Phe	Leu	Phe	Gln	His	Trp	Arg	Thr	195	200	205	
Gly	Tyr	Ser	Lys	Ser	Ser	His	Pro	Leu	Ile	Arg	Ser	Leu	Phe	His	Gly	210	215	220	
Phe	Leu	Phe	Met	Ile	Phe	Gln	Ile	Gly	Val	Leu	Gly	Phe	Gly	Pro	Thr	225	230	235	240
Tyr	Val	Val	Leu	Ala	Tyr	Thr	Leu	Pro	Pro	Ala	Ser	Arg	Phe	Ile	Ile	245	250	255	
Ile	Phe	Glu	Gln	Ile	Arg	Phe	Val	Met	Lys	Ala	His	Ser	Phe	Val	Arg	260	265	270	
Glu	Asn	Val	Pro	Arg	Val	Leu	Asn	Ser	Ala	Lys	Glu	Lys	Ser	Ser	Thr	275	280	285	
Val	Pro	Ile	Pro	Thr	Val	Asn	Gln	Tyr	Leu	Tyr	Phe	Leu	Phe	Ala	Pro	290	295	300	

-60-

Thr Leu Ile Tyr Arg Asp Ser Tyr Pro Arg Asn Pro Thr Val Arg Trp
 305 310 315 320
 Gly Tyr Val Ala Met Lys Phe Ala Gln Val Phe Gly Cys Phe Phe Tyr
 325 330 335
 Val Tyr Tyr Ile Phe Glu Arg Leu Cys Ala Pro Leu Phe Arg Asn Ile
 340 345 350
 Lys Gln Glu Pro Phe Ser Ala Arg Val Leu Val Leu Cys Val Phe Asn
 355 360 365
 Ser Ile Leu Pro Gly Val Leu Ile Leu Phe Leu Thr Phe Phe Ala Phe
 370 375 380
 Leu His Cys Trp Leu Asn Ala Phe Ala Glu Met Leu Arg Phe Gly Asp
 385 390 395 400
 Arg Met Phe Tyr Lys Asp Trp Trp Asn Ser Thr Ser Tyr Ser Asn Tyr
 405 410 415
 Tyr Arg Thr Trp Asn Val Val Val His Asp Trp Leu Tyr Tyr Tyr Ala
 420 425 430
 Tyr Lys Asp Phe Leu Trp Phe Phe Ser Lys Arg Phe Lys Ser Ala Ala
 435 440 445
 Met Leu Ala Val Phe Ala Val Ser Ala Val Val His Glu Tyr Ala Leu
 450 455 460
 Ala Val Cys Leu Ser Phe Phe Tyr Pro Val Leu Phe Val Leu Phe Met
 465 470 475 480
 Phe Phe Gly Met Ala Phe Asn Phe Ile Val Asn Asp Ser Arg Lys Lys
 485 490 495
 Pro Ile Trp Asn Val Leu Met Trp Thr Ser Leu Phe Leu Gly Asn Gly
 500 505 510
 Val Leu Leu Cys Phe Tyr Ser Gln Glu Trp Tyr Ala Arg Arg His Cys
 515 520 525
 Pro Leu Lys Asn Pro Thr Phe Leu Asp Tyr Val Arg Pro Arg Ser Trp
 530 535 540
 Thr Cys Arg Tyr Val Phe
 545 550

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2601 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

-61-

GCCCTCCAGC TCTCTACTAA GACCGGTCGC AAGCATGCTG GGCGATATAT CCAAACCACA	60
CCACACATGG TCTCCCTCCT GCGTCAAAAT CTCCCAGAC AGTCCGGACC CGCACCCGAT	120
ATCCAGAATG AAAGTGCACG GCTGCAGATT CAAAAGCTCC AACGCCCTCA GCGTCATCTT	180
CGCCTGGATA TGCTGCACTC TGGTCGAACC CGTGTACTTG TGTGCTTCGC TATCATTATA	240
GAAAATCTCC GGTGGTGCCA ACTCCTCAGG ACGTGACATT ATTTCTTCTC TGATATATTT	300
CCTGTGTTTC CGTACCGCAC CTTTTTAGCA CTACTTTTTT ACTATGCTCT TCTTCTTCTG	360
CTTCTTCTGC TTTTTTCCTC TTTATCACAC TATGTATGTG CTGCTCATCT CTTCTTTTTA	420
TCGATAAAAT TGAAAAATGT GAGATGGTGT AGAGTGAAAA AAAAAAAAAA ATCTGGCTTG	480
GCCATCAAAT ACCCGGCCGT GGTTGGACTC GTTTAGCGAA CAATAGCACC CAGCAGACCC	540
TGGCAACATG CGGATGATAT AAGAAGGACG AGCGTGGTGG AGGAAAGGGG CGCCATTGGC	600
ACACTCACGC AGGTGGTTGT TCAGCACGGC TTGCAGCAAG AGCGCCAAAA CAGATTGCAA	660
GAATGACGGA GACTAAGGAT TTGTTGCAAG ACGAAGAGTT TCTTAAGATC CGCAGACTCA	720
ATTCCGCAGA AGCCAACAAA CGGCATTCCG TCACGTACGA TAACGTGATC CTGCCACAGG	780
AGTCCATGGA GGTTTCGCCA CGGTCGTCTA CCACGTCGCT GGTGGAGCCA GTGGAGTCGA	840
CTGAAGGAGT GGAGTCGACT GAGGCGGAAC GTGTGGCAGG GAAGCAGGAG CAGGAGGAGG	900
AGTACCCTGT GGACGCCAC ATGCAAAAGT ACCTTTCACA CCTGAAGAGC AAGTCTCGGT	960
CGAGGTTCCA CCGAAAGGAT GCTAGCAAGT ATGTGTCGTT TTTTGGGGAC GTGAGTTTTG	1020
ATCCTCGCCC CACGCTCCTG GACAGCGCCA TCAACGTGCC CTTCCAGACG ACTTTCAAAG	1080
GTCCGGTGCT GGAGAAACAG CTCAAAAATT TACAGTTGAC AAAGACCAAG ACCAAGGCCA	1140
CGGTGAAGAC TACGGTGAAG ACTACGGAGA AAACGGACAA GGCAGATGCC CCCCAGGAG	1200
AAAAACTGGA GTCGAACTTT TCAGGGATCT ACGTGTCGC ATGGATGTTT TGGGCTGGA	1260
TAGCCATCAG GTGCTGCACA GATTACTATG CGTCGTACGG CAGTGCATGG AATAAGCTGG	1320
AAATCGTGCA GTACATGACA ACGGACTTGT TCACGATCGC AATGTTGGAC TTGGCAATGT	1380
TCCTGTGCAC TTTCTTCGTG GTTTTCGTGC ACTGGCTGGT GAAAAAGCGG ATCATCAACT	1440
GGAAGTGGAC TGGGTTTCGTT GCAGTGAGCA TCTTCGAGTT GGCTTTCATC CCCGTGACGT	1500
TCCCCATTTA CGTCTACTAC TTTGATTTCA ACTGGGTCAC GAGAATCTTC CTGTTCTTGC	1560
ACTCCGTGGT GTTTGTTATG AAGAGCCACT CGTTTGCCTT TTACAACGGG TATCTTTGGG	1620
ACATAAAGCA GGAAGTCGAG TACTCTTCCA AACAGTTGCA AAAATACAAG GAATCTTTGT	1680
CCCCAGAGAC CCGCGAGATT CTGCAAAAAA GTTGCAGACTT TTGCCTTTTC GAATTGAACT	1740
ACCAGACCAA GGATAACGAC TTCCCCAACA ACATCAGTTG CAGCAATTTT TTCATGTTCT	1800
GTTTGTTCCC CGTCCTCGTG TACCAGATCA ACTACCCAAG AACGTCGCGC ATCAGATGGA	1860
GGTATGTGTT GGAGAAGGTG TCGCCATCA TTGGCACCAT CTTCTCATG ATGGTCACGG	1920

-62-

CACAGTTCTT CATGCACCCG GTGGCCATGC GCTGTATCCA GTTCCACAAC ACGCCACCT 1980
TCGGCGGCTG GATCCCCGCC ACGCAAGAGT GGTTCCACCT GCTCTTCGAC ATGATTCCGG 2040
GCTTCACTGT TCTGTACATG CTCACGTTTT ACATGATATG GGACGCTTTA TTGAATTGCG 2100
TGGCGGAGTT GACCAGGTTT GCGGACAGAT ATTTCTACGG CGACTGGTGG AATTGCGTTT 2160
CGTTTGAAGA GTTTAGCAGA ATCTGGAACG TCCCCGTTCA CAAATTTTTA CTAAGACACG 2220
TGTACCACAG CTCCATGGGC GCATTGCATT TGAGCAAGAG CCAAGCTACA TTATTTACTT 2280
TTTTCTTGAG TGCCGTGTTT CACGAAATGG CCATGTTTCGC CATTTTCAGA AGGGTTAGAG 2340
GATATCTGTT CATGTTCCAA CTGTCGCAGT TTGTGTGGAC TGCTTTGAGC AACACCAAGT 2400
TTCTACCGGC AAGACCGCAG TTGTCCAACG TTGTCTTTTC GTTTGGTGTC GTTTCAGGGC 2460
CCAGTATCAT TATGACGTTG TACCTGACCT TATGAACTGC CACCATACCA CGTGTGTCCC 2520
TCGCAAGCCC TTGATAGATA TACAATAGGG AATGGGCGTC CGTCCACCGT GGTCAAAGAC 2580
AGGGGCAAAG AGCTCCTAGG T

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 610 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Thr Glu Thr Lys Asp Leu Leu Gln Asp Glu Glu Phe Leu Lys Ile
1 5 10 15
Arg Arg Leu Asn Ser Ala Glu Ala Asn Lys Arg His Ser Val Thr Tyr
20 25 30
Asp Asn Val Ile Leu Pro Gln Glu Ser Met Glu Val Ser Pro Arg Ser
35 40 45
Ser Thr Thr Ser Leu Val Glu Pro Val Glu Ser Thr Glu Gly Val Glu
50 55 60
Ser Thr Glu Ala Glu Arg Val Ala Gly Lys Gln Glu Gln Glu Glu Glu
65 70 75 80
Tyr Pro Val Asp Ala His Met Gln Lys Tyr Leu Ser His Leu Lys Ser
85 90 95
Lys Ser Arg Ser Arg Phe His Arg Lys Asp Ala Ser Lys Tyr Val Ser
100 105 110
Phe Phe Gly Asp Val Ser Phe Asp Pro Arg Pro Thr Leu Leu Asp Ser
115 120 125

-63-

Ala Ile Asn Val Pro Phe Gln Thr Thr Phe Lys Gly Pro Val Leu Glu
 130 135 140
 Lys Gln Leu Lys Asn Leu Gln Leu Thr Lys Thr Lys Thr Lys Ala Thr
 145 150 155 160
 Val Lys Thr Thr Val Lys Thr Thr Glu Lys Thr Asp Lys Ala Asp Ala
 165 170 175
 Pro Pro Gly Glu Lys Leu Glu Ser Asn Phe Ser Gly Ile Tyr Val Phe
 180 185 190
 Ala Trp Met Phe Leu Gly Trp Ile Ala Ile Arg Cys Cys Thr Asp Tyr
 195 200 205
 Tyr Ala Ser Tyr Gly Ser Ala Trp Asn Lys Leu Glu Ile Val Gln Tyr
 210 215 220
 Met Thr Thr Asp Leu Phe Thr Ile Ala Met Leu Asp Leu Ala Met Phe
 225 230 235 240
 Leu Cys Thr Phe Phe Val Val Phe Val His Trp Leu Val Lys Lys Arg
 245 250 255
 Ile Ile Asn Trp Lys Trp Thr Gly Phe Val Ala Val Ser Ile Phe Glu
 260 265 270
 Leu Ala Phe Ile Pro Val Thr Phe Pro Ile Tyr Val Tyr Tyr Phe Asp
 275 280 285
 Phe Asn Trp Val Thr Arg Ile Phe Leu Phe Leu His Ser Val Val Phe
 290 295 300
 Val Met Lys Ser His Ser Phe Ala Phe Tyr Asn Gly Tyr Leu Trp Asp
 305 310 315 320
 Ile Lys Gln Glu Leu Glu Tyr Ser Ser Lys Gln Leu Gln Lys Tyr Lys
 325 330 335
 Glu Ser Leu Ser Pro Glu Thr Arg Glu Ile Leu Gln Lys Ser Cys Asp
 340 345 350
 Phe Cys Leu Phe Glu Leu Asn Tyr Gln Thr Lys Asp Asn Asp Phe Pro
 355 360 365
 Asn Asn Ile Ser Cys Ser Asn Phe Phe Met Phe Cys Leu Phe Pro Val
 370 375 380
 Leu Val Tyr Gln Ile Asn Tyr Pro Arg Thr Ser Arg Ile Arg Trp Arg
 385 390 395 400
 Tyr Val Leu Glu Lys Val Cys Ala Ile Ile Gly Thr Ile Phe Leu Met
 405 410 415
 Met Val Thr Ala Gln Phe Phe Met His Pro Val Ala Met Arg Cys Ile
 420 425 430
 Gln Phe His Asn Thr Pro Thr Phe Gly Gly Trp Ile Pro Ala Thr Gln
 435 440 445
 Glu Trp Phe His Leu Leu Phe Asp Met Ile Pro Gly Phe Thr Val Leu
 450 455 460

- 64 -

Tyr Met Leu Thr Phe Tyr Met Ile Trp Asp Ala Leu Leu Asn Cys Val
 465 470 475 480
 Ala Glu Leu Thr Arg Phe Ala Asp Arg Tyr Phe Tyr Gly Asp Trp Trp
 485 490 495
 Asn Cys Val Ser Phe Glu Glu Phe Ser Arg Ile Trp Asn Val Pro Val
 500 505 510
 His Lys Phe Leu Leu Arg His Val Tyr His Ser Ser Met Gly Ala Leu
 515 520 525
 His Leu Ser Lys Ser Gln Ala Thr Leu Phe Thr Phe Phe Leu Ser Ala
 530 535 540
 Val Phe His Glu Met Ala Met Phe Ala Ile Phe Arg Arg Val Arg Gly
 545 550 555 560
 Tyr Leu Phe Met Phe Gln Leu Ser Gln Phe Val Trp Thr Ala Leu Ser
 565 570 575
 Asn Thr Lys Phe Leu Arg Ala Arg Pro Gln Leu Ser Asn Val Val Phe
 580 585 590
 Ser Phe Gly Val Cys Ser Gly Pro Ser Ile Ile Met Thr Leu Tyr Leu
 595 600 605
 Thr Leu
 610

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 2421 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TATAAAATTC CTTTCATCAA TACATCTATA TATTCGAATA TATAGATAAA CCAATACAAA	60
AACATACTGA AATTTTTTGA AAACAACATA AACTATTCAT TGCAGTTACA CGTGAATGCT	120
AAACTTTATA TCGCTCTGT CGGTCCCGCG GAGTTAATCAT TTAACGGCTT CTCGCGCAAT	180
AACCGGAAAA ATTCCAACAG TTTCTTTGTA ATATTATTAA GCCTTCTTTT TTCCCGGAAT	240
CTATAAGAGG GGACGAAAAT TAGCCGCTAT TAATTCTGGT ATTGCCACCT AGACAAGAAG	300
TAAACAGACA CATTACGTTA GCAAAAGCAA CAATAACAAA CACAACCATG GACAAGAAGA	360
AGGATCTACT GGAGAACGAA CAATTTCTCC GCATCCAAAA GCTCAACGCT GCCGATGCGG	420
GCAAAAGACA ATCTATAACA GTGGACGACG AGGGCGAACT ATATGGGTTA GACACCTCCG	480
GCAACTCACC AGCCAATGAA CACACAGCTA CCACAATTAC ACAGAATCAC AGCGTGGTGG	540

-65-

CCTCAAACGG AGACGTCGCA TTCATCCCAG GAACTGCTAC CGAAGGCAAT ACAGAGATTG	600
TAACTGAAGA AGTGATTGAG ACCGATGATA ACATGTTCAA GACCCATGTG AAGACTTTAA	660
GCTCCAAAGA GAAGGCACGG TATAGGCAAG GGTCTCTAA CTTTATATCG TATTTGATG	720
ATATGTCATT TGAACACAGG CCCAGTATAT TAGATGGGTC AGTTAACGAG CCCTTCAAGA	780
CCAAATTCGT GGGACCTACT TTAGAAAAGG AGATCAGAAG AGGGGAGAAA GAGCTAATGG	840
CCATGCGCAA AAATTTACAC CACCGCAAGT CCTCCCCAGA TGCTGTCGAC TCAGTAGGGA	900
AAAATGATGG CGCCGCCCCA ACTACTGTTC CAACTGCCGC CACCTCAGAA ACGGTGGTCA	960
CCGTTGAAAC CACCATAATT TCATCCAATT TCTCCGGGTT GTACGTGGCG TTTTGGATGG	1020
CTATTGCATT TGGTGCTGTC AAGGCTTTAA TAGACTATTA TTACCAGCAT AATGGTAGCT	1080
TCAAGGATTC GGAGATCTTG AAATTTATGA CTACGAATTT GTTCACTGTG GCATCCGTAG	1140
ATCTTTTGAT GTATTTGAGC ACTTATTTTG TCGTTGGAAT ACAATACTTA TGCAAGTGGG	1200
GGGTCTTGAA ATGGGGCACT ACCGGCTGGA TCTTCACCTC AATTTACGAG TTTTGTGTTG	1260
TTATCTTCTA CATGTATTTA ACAGAAAACA TCCTAAAACT ACACTGGCTG TCCAAGATCT	1320
TCCTTTTTTT GCATTCTTTA GTTTTATTGA TGAAAATGCA TTCTTTCGCC TTCTACAATG	1380
GCTATCTATG GGGTATAAAG GAAGAACTAC AATTTTCCAA AAGCGCTCTT GCCAAATACA	1440
AGGATTCTAT AAATGATCCA AAAGTTATTG GTGCTCTTGA GAAAAGCTGT GAGTTTTGTA	1500
GTTTTGAATT GAGCTCTCAG TCTTTAAGCG ACCAACTCA AAAATTCCCC AACAATATCA	1560
GTGCAAAAAG CTTTTTTTGG TTCACCATGT TTCCAACCTT AATTTACCAA ATTGAATATC	1620
CAAGAACTAA GGAAATCAGA TGGAGCTACG TATTAGAAAA GATCTGCGCC ATCTTCGGTA	1680
CCATTTTCTT AATGATGATA GATGCTCAA TCTTGATGTA TCCTGTAGCA ATGAGAGCAT	1740
TGGCTGTGCG CAATTCTGAA TGGACTGGTA TATTGGATAG ATTATTGAAA TGGGTGGAT	1800
TGCTCGTTGA TATCGTCCCA GGGTTTATCG TGATGTACAT CTTGGACTTC TATTTGATTT	1860
GGGATGCCAT TTTGAACTGT GTGGCTGAAT TGACAAGATT TGGCGACAGA TATTTCTACG	1920
GTGACTGGTG GAATTGTGTT AGTTGGGCAG ACTTCAGTAG AATTTGGAAC ATCCCAGTGC	1980
ATAAGTTTTT GTTAAGACAT GTTTACCATA GTTCAATGAG TTCATTCAAA TTGAACAAGA	2040
GTCAAGCAAC TTTGATGACC TTTTCTTAA GTTCCGTCGT TCATGAATTA GCAATGTACG	2100
TTATCTTCAA GAAATTGAGG TTTTACTTGT TCTTCTTCCA AATGCTGCAA ATGCCATTAG	2160
TAGCTTTAAC AAATACTAAA TTCATGAGGA ACAGAACCAT AATCGGAAAT GTTATTTTCT	2220
GGCTCGGTAT CTGCATGGGA CCAAGTGTCA TGTGTACGTT GACTTGACA TTCTAAGGCA	2280
TCCTGCAACT GTTCTGTGGA GCTATTAAAT CTTTATAGTA AATTTTTTTT TACTTTTTTT	2340
TTTTTTTTTT TTTTTTTT TATTTTACAA GCGTCTATAT ATTTTCTATT ATAGAATATT	2400
GTCATTTATT ACATTGGTTC A	

-66-

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 642 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Asp Lys Lys Lys Asp Leu Leu Glu Asn Glu Gln Phe Leu Arg Ile
1      5      10      15
Gln Lys Leu Asn Ala Ala Asp Ala Gly Lys Arg Gln Ser Ile Thr Val
20      25      30
Asp Asp Glu Gly Glu Leu Tyr Gly Leu Asp Thr Ser Gly Asn Ser Pro
35      40      45
Ala Asn Glu His Thr Ala Thr Thr Ile Thr Gln Asn His Ser Val Val
50      55      60
Ala Ser Asn Gly Asp Val Ala Phe Ile Pro Gly Thr Ala Thr Glu Gly
65      70      75      80
Asn Thr Glu Ile Val Thr Glu Glu Val Ile Glu Thr Asp Asp Asn Met
85      90      95
Phe Lys Thr His Val Lys Thr Leu Ser Ser Lys Glu Lys Ala Arg Tyr
100     105     110
Arg Gln Gly Ser Ser Asn Phe Ile Ser Tyr Phe Asp Asp Met Ser Phe
115     120     125
Glu His Arg Pro Ser Ile Leu Asp Gly Ser Val Asn Glu Pro Phe Lys
130     135     140
Thr Lys Phe Val Gly Pro Thr Leu Glu Lys Glu Ile Arg Arg Arg Glu
145     150     155     160
Lys Glu Leu Met Ala Met Arg Lys Asn Leu His His Arg Lys Ser Ser
165     170     175
Pro Asp Ala Val Asp Ser Val Gly Lys Asn Asp Gly Ala Ala Pro Thr
180     185     190
Thr Val Pro Thr Ala Ala Thr Ser Glu Thr Val Val Thr Val Glu Thr
195     200     205
Thr Ile Ile Ser Ser Asn Phe Ser Gly Leu Tyr Val Ala Phe Trp Met
210     215     220
Ala Ile Ala Phe Gly Ala Val Lys Ala Leu Ile Asp Tyr Tyr Tyr Gln
225     230     235     240
His Asn Gly Ser Phe Lys Asp Ser Glu Ile Leu Lys Phe Met Thr Thr
245     250     255

```

-67-

Asn Leu Phe Thr Val Ala Ser Val Asp Leu Leu Met Tyr Leu Ser Thr
 260 265 270
 Tyr Phe Val Val Gly Ile Gln Tyr Leu Cys Lys Trp Gly Val Leu Lys
 275 280 285
 Trp Gly Thr Thr Gly Trp Ile Phe Thr Ser Ile Tyr Glu Phe Leu Phe
 290 295 300
 Val Ile Phe Tyr Met Tyr Leu Thr Glu Asn Ile Leu Lys Leu His Trp
 305 310 315 320
 Leu Ser Lys Ile Phe Leu Phe Leu His Ser Leu Val Leu Leu Met Lys
 325 330 335
 Met His Ser Phe Ala Phe Tyr Asn Gly Tyr Leu Trp Gly Ile Lys Glu
 340 345 350
 Glu Leu Gln Phe Ser Lys Ser Ala Leu Ala Lys Tyr Lys Asp Ser Ile
 355 360 365
 Asn Asp Pro Lys Val Ile Gly Ala Leu Glu Lys Ser Cys Glu Phe Cys
 370 375 380
 Ser Phe Glu Leu Ser Ser Gln Ser Leu Ser Asp Gln Thr Gln Lys Phe
 385 390 395 400
 Pro Asn Asn Ile Ser Ala Lys Ser Phe Phe Trp Phe Thr Met Phe Pro
 405 410 415
 Thr Leu Ile Tyr Gln Ile Glu Tyr Pro Arg Thr Lys Glu Ile Arg Trp
 420 425 430
 Ser Tyr Val Leu Glu Lys Ile Cys Ala Ile Phe Gly Thr Ile Phe Leu
 435 440 445
 Met Met Ile Asp Ala Gln Ile Leu Met Tyr Pro Val Ala Met Arg Ala
 450 455 460
 Leu Ala Val Arg Asn Ser Glu Trp Thr Gly Ile Leu Asp Arg Leu Leu
 465 470 475 480
 Lys Trp Val Gly Leu Leu Val Asp Ile Val Pro Gly Phe Ile Val Met
 485 490 495
 Tyr Ile Leu Asp Phe Tyr Leu Ile Trp Asp Ala Ile Leu Asn Cys Val
 500 505 510
 Ala Glu Leu Thr Arg Phe Gly Asp Arg Tyr Phe Tyr Gly Asp Trp Trp
 515 520 525
 Asn Cys Val Ser Trp Ala Asp Phe Ser Arg Ile Trp Asn Ile Pro Val
 530 535 540
 His Lys Phe Leu Leu Arg His Val Tyr His Ser Ser Met Ser Ser Phe
 545 550 555 560
 Lys Leu Asn Lys Ser Gln Ala Thr Leu Met Thr Phe Phe Leu Ser Ser
 565 570 575
 Val Val His Glu Leu Ala Met Tyr Val Ile Phe Lys Lys Leu Arg Phe
 580 585 590

-68-

Tyr Leu Phe Phe Phe Gln Met Leu Gln Met Pro Leu Val Ala Leu Thr
 595 600 605

Asn Thr Lys Phe Met Arg Asn Arg Thr Ile Ile Gly Asn Val Ile Phe
 610 615 620

Trp Leu Gly Ile Cys Met Gly Pro Ser Val Met Cys Thr Leu Tyr Leu
 625 630 635 640

Thr Phe

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 983 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

ATGGAGCTCA ACTTTCCCCG CTCTCCCCGC ATCCGGAAGC GCTTTCTGCT GCGACGGATC      60
CTTGAGATGC TGTTCTTCAC CCAGCTCCAG GTGGGGCTGA TCCAGCAGTG GATGGTCCCC      120
ACCATCCAGA ACTCCATGAA GCCCTTCAAG GACATGGACT ACTCACGCAT CATCGAGCGC      180
CTCCTGAAGC TGGCGGTCCC CAATCACCTC ATCTGGCTCA TCTTCTTCTA CTGGCTCTTC      240
CACTCCTGCC TGAATGCCGT GGCTGAGCTC ATGCAGTTTG GAGACCGGGA GTTCTACCGG      300
GACTGGTGGA ACTCCGAGTC TGTCACCTAC TTCTGGCAGA ACTGGAACAT CCCTGTGCAC      360
AAGTGGTGCA TCAGACACTT CTACAAGCCC ATGCTTCGAC GGGGCAGCAG CAAGTGGATG      420
GCCAGGACAG GGGTGTTCCCT GGCCTCGGCT TTCTTCCACG AGTACCTGGT GAGCGTCCCT      480
CTGCGAATGT TCCGCCTCTG GGCTTTCACG GGCATGATGG CTCAGATCCC ACTGGCCTGG      540
TTCGTGGGCC GCTTTTTCCTA GGGCAACTAT GGCAACGCAG CTGTGTGGCT GTCGCTCATC      600
ATCGGACAGC CAATAGCCGT CCTCATGTAC GTCCACGAAC TACTACGTGC TCAACTATGA      660
GGCCCCAGCG GCAGAGGCCT GAGCTGCACC TGAGGGCCTG GCTTCTCACT GCCACCTCAA      720
ACCCGCTGCC AGAGCCCACC TCTCCTCCTA GGCCTCGAGT GCTGGGGATG GGCCTGGCTG      780
CACAGCATCC TCCTCTGGTC CCAGGGAGGC CTCTCTGCCC TATGGGGCTC TGTCTTGCAC      840
CCCTCAGGGA TGGCGACAGC AGGCCAGACA CAGTCTGATG CCAGCTGGGA GTCTTGCTGA      900
CCCTGCCCCG GGTCCGAGGG TGTCATAAAA GTGCTGTCCA GTGAGAAAAA GAAAAAAAAA      960
AAAAAAAAAA ATTCTGCGGC CGC

```

(2) INFORMATION FOR SEQ ID NO:8:

-69-

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 219 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```

Met Glu Leu Asn Phe Pro Arg Ser Pro Arg Ile Arg Lys Arg Phe Leu
 1             5             10             15
Leu Arg Arg Ile Leu Glu Met Leu Phe Phe Thr Gln Leu Gln Val Gly
      20             25             30
Leu Ile Gln Gln Trp Met Val Pro Thr Ile Gln Asn Ser Met Lys Pro
      35             40             45
Phe Lys Asp Met Asp Tyr Ser Arg Ile Ile Glu Arg Leu Leu Lys Leu
      50             55             60
Ala Val Pro Asn His Leu Ile Trp Leu Ile Phe Phe Tyr Trp Leu Phe
      65             70             75             80
His Ser Cys Leu Asn Ala Val Ala Glu Leu Met Gln Phe Gly Asp Arg
      85             90             95
Glu Phe Tyr Arg Asp Trp Trp Asn Ser Glu Ser Val Thr Tyr Phe Trp
      100             105             110
Gln Asn Trp Asn Ile Pro Val His Lys Trp Cys Ile Arg His Phe Tyr
      115             120             125
Lys Pro Met Leu Arg Arg Gly Ser Ser Lys Trp Met Ala Arg Thr Gly
      130             135             140
Val Phe Leu Ala Ser Ala Phe Phe His Glu Tyr Leu Val Ser Val Pro
      145             150             155             160
Leu Arg Met Phe Arg Leu Trp Ala Phe Thr Gly Met Met Ala Gln Ile
      165             170             175
Pro Leu Ala Trp Phe Val Gly Arg Phe Phe Gln Gly Asn Tyr Gly Asn
      180             185             190
Ala Ala Val Trp Leu Ser Leu Ile Ile Gly Gln Pro Ile Ala Val Leu
      195             200             205
Met Tyr Val His Glu Leu Leu Arg Ala Gln Leu
      210             215

```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 455 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-70-

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```

ATGTTGAACT TCATGATGCA TGACCAGCGC ACCGGCCCCG CATGGAACGT GCTGATGTGG      60
ACCATGCTGT TTCTAGGCCA GGAATCCAG GTCAGCCTGT ACTGCCAGGA GTGGTACGCA      120
CGGACGCACT GCCCCCTTACC CCAGGCAACT TTCTGGGGGC TGGTGACACC TCGATCTTGG      180
TCCTGCCATA CCTAGAGGTC GGGACAGACG ACGCTACCTG CCCAGACACC ACCAAGTTCT      240
CTGCCTGCAA AACCTGGGGA CCAGGACTTC CTGTCTTGCA TTCCCAAATT TGGGTTCTTG      300
AGTCGAGGCA ACCTTGCACA CAAGACCCCA CCAAGGGATT GTTGCAAGGG ATTAGATTTT      360
GCAGATTTGT TGGGTAATGA TTCAACGACT CAGCTGGGGG TTGACCAGGG TTGATTTTTC      420
AATCCTTTTC CCCTGGGTTT GGGTTACAGG TTTTTC      455

```

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 64 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```

Met Leu Asn Phe Met Met His Asp Gln Arg Thr Gly Pro Ala Trp Asn
1           5           10           15
Val Leu Met Trp Thr Met Leu Phe Leu Gly Gln Gly Ile Gln Val Ser
20           25           30
Leu Tyr Cys Gln Glu Trp Tyr Ala Arg Thr His Cys Pro Leu Pro Gln
35           40           45
Ala Thr Phe Trp Gly Leu Val Thr Pro Arg Ser Trp Ser Cys His Thr
50           55           60

```

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 517 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

-71-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```

ATGGACAACG CGGGGTCTGA TACGACTCAC TATAGGGAAT TTGGCCCTCG AGCAGTAGAT      60
TCGGCACGAT GGGCAGGAGG ACTCCATCAT GTTCCTCAAG CTTTATTCCT ACCGGGATGT      120
CAACCTGTGG TGCCGCCAGC GAAGGGTCAA GGCCAAAGCT GTCTCTACAG GGAAGAAGGT      180
CAGTGGGGCT GCTGCGAGCA AGCTGTGAGC TATCCAGACA ACCTGACCTA CCGAGATCTC      240
GATTACTTCA TCTTTGCTCC TACTTTGTGT TATGAACTCA ACTTTCCTCG GTCCCCCGA      300
ATACGAGAGC GCTTCTGCT ACGACGAGTT CTTGAGATGC TCTTTTTTAC CCAGCTTCAA      360
GTGGGGCTGA TCCAACAGTG GATGGTCCCT ACTATCCAGA ACTCCATGGA AGCCCTTTCA      420
AGAGCTTCTG GCAGTTTGG AGACGCGAG TTCTACAGAG ATTGGTGGAA TGCTGAGTCT      480
GTCACCGACT TTTGGCAGAA CTGGAATATC CCCGTGG      517

```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 172 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```

Met Asp Asn Ala Gly Ser Asp Thr Thr His Tyr Arg Glu Phe Gly Pro
1          5          10          15
Arg Ala Val Asp Ser Ala Arg Trp Ala Arg Gly Leu His His Val Pro
20          25          30
Gln Ala Leu Phe Leu Pro Gly Cys Gln Pro Val Val Pro Pro Ala Lys
35          40          45
Gly Gln Gly Gln Ser Cys Leu Tyr Arg Glu Glu Gly Gln Trp Gly Cys
50          55          60
Cys Glu Gln Ala Val Ser Tyr Pro Asp Asn Leu Thr Tyr Arg Asp Leu
65          70          75          80
Asp Tyr Phe Ile Phe Ala Pro Thr Leu Cys Tyr Glu Leu Asn Phe Pro
85          90          95
Arg Ser Pro Arg Ile Arg Glu Arg Phe Leu Leu Arg Arg Val Leu Glu
100         105         110
Met Leu Phe Phe Thr Gln Leu Gln Val Gly Leu Ile Gln Gln Trp Met
115         120         125
Val Pro Thr Ile Gln Asn Ser Met Glu Ala Leu Ser Arg Ala Ser Gly
130         135         140
Ser Phe Gly Asp Arg Glu Phe Tyr Arg Asp Trp Trp Asn Ala Glu Ser

```

-72-

145		150		155		160
Val Thr Asp Phe Trp Gln Asn Trp Asn Ile Pro Val						
		165		170		

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 366 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met	Lys	Asp	Leu	Leu	Glu	Phe	Leu	Lys	Ile	Arg	Leu	Asn	Ala	Asp	Ala
1				5					10					15	
Lys	Arg	Ser	Thr	Asp	Ser	Pro	Thr	Val	Ser	Glu	Val	Glu	Arg	Gly	Lys
			20					25					30		
Gln	Glu	Ile	Glu	Ala	His	Lys	Ser	Lys	Lys	Arg	Phe	Arg	Ser	Phe	Ser
		35					40					45			
Phe	Phe	Asp	Ser	Phe	Glu	Arg	Pro	Ser	Leu	Leu	Asp	Gly	Asn	Pro	Phe
	50					55					60				
Thr	Thr	Phe	Gly	Pro	Val	Leu	Glu	Lys	Glu	Lys	Asn	Leu	His	Lys	Lys
65					70				75					80	
Lys	Thr	Thr	Val	Thr	Asp	Val	Ser	Asn	Phe	Ser	Gly	Ile	Tyr	Val	Phe
				85					90					95	
Trp	Met	Leu	Ala	Leu	Asp	Tyr	Tyr	Gly	Glu	Ile	Leu	Tyr	Met	Thr	Thr
			100					105					110		
Leu	Phe	Thr	Val	Ala	Asp	Leu	Met	Phe	Leu	Ser	Thr	Phe	Phe	Val	Val
		115					120					125			
Leu	Lys	Trp	Thr	Gly	Ile	Ser	Ile	Glu	Phe	Leu	Phe	Ile	Phe	Leu	Trp
	130					135					140				
Ser	Arg	Ile	Phe	Leu	Phe	Leu	His	Ser	Val	Phe	Val	Met	Lys	His	Ser
145					150					155				160	
Phe	Ala	Phe	Tyr	Asn	Gly	Tyr	Leu	Trp	Ile	Lys	Glu	Glu	Leu	Ser	Leu
				165					170					175	
Lys	Tyr	Lys	Glu	Ser	Ser	Pro	Leu	Gln	Lys	Ser	Cys	Phe	Cys	Phe	Glu
			180					185					190		
Leu	Gln	Phe	Pro	Asn	Asn	Ile	Ser	Phe	Phe	Phe	Phe	Pro	Thr	Leu	Ile
		195					200					205			
Tyr	Gln	Ile	Tyr	Pro	Arg	Thr	Ile	Arg	Trp	Tyr	Val	Leu	Glu	Lys	Cys
	210					215					220				
Ala	Ile	Phe	Gly	Thr	Ile	Phe	Leu	Met	Met	Ala	Gln	Met	Pro	Val	Ala

-73-

225		230		235		240									
Met	Arg	Asn	Phe	Trp	Gln	Leu	Leu	Asp	Ile	Pro	Gly	Phe	Val	Leu	Tyr
				245					250					255	
Leu	Thr	Phe	Tyr	Ile	Trp	Asp	Ala	Leu	Asn	Cys	Val	Ala	Glu	Leu	Thr
			260					265					270		
Arg	Phe	Gly	Asp	Arg	Tyr	Phe	Tyr	Gly	Asp	Trp	Trp	Asn	Cys	Val	Ser
		275					280					285			
Phe	Ser	Arg	Ile	Trp	Asn	Val	Pro	Val	His	Lys	Phe	Leu	Leu	Arg	His
	290					295					300				
Val	Tyr	His	Ser	Ser	Met	Phe	Lys	Leu	Lys	Ser	Gln	Ala	Thr	Leu	Thr
305					310					315				320	
Phe	Phe	Leu	Ser	Ala	Val	Val	His	Glu	Ala	Met	Val	Ile	Phe	Arg	Tyr
				325					330					335	
Leu	Phe	Phe	Gln	Gln	Met	Ala	Leu	Asn	Thr	Lys	Phe	Arg	Arg	Ile	Asn
			340					345					350		
Val	Phe	Trp	Gly	Cys	Gly	Pro	Ser	Val	Thr	Leu	Tyr	Leu	Thr		
		355					360					365			

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 96 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Amino Acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Pro	Asn	His	Leu	Ile	Trp	Leu	Ile	Phe	Phe	Tyr	Trp	Leu	Phe	His	Ser
1			5					10						15	
Cys	Leu	Asn	Ala	Val	Ala	Glu	Leu	Met	Gln	Phe	Gly	Asp	Arg	Glu	Phe
			20					25					30		
Tyr	Arg	Asp	Trp	Trp	Asn	Ser	Glu	Ser	Val	Thr	Tyr	Phe	Trp	Gln	Asn
		35				40						45			
Trp	Lys	Ile	Pro	Val	His	Lys	Trp	Cys	Ile	Arg	His	Phe	Tyr	Lys	Pro
	50					55					60				
Met	Leu	Arg	Arg	Gly	Ser	Ser	Lys	Trp	Met	Ala	Arg	Asp	Arg	Gly	Val
65				70					75					80	
Pro	Gly	Pro	Ser	Ala	Phe	Phe	His	Val	Val	Thr	Trp	Val	Ser	Val	Pro
				85				90					95		

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:

-74-

- (A) LENGTH: 91 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAGGGGACGA AAATTAGCCG CTATTAATTC TGGTATTGCC ACCTAGACAA GAAGTAAACA 60
GACACAGATG CAAGAGTTCG AATCTCTTAG C 91

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 76 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CTATAAAGAT TTAATAGCTC CACAGAACAG TTGCAGGATG CCTTAGGGTC GACTACGTCG 60
TAAGGCCGTT TCTGAC 76

(2) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CATTGCAGTT ACACGTGAAT GC 22

(2) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

-75-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TAGCTCCACA GAACAGTTGC AGG

23

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CTCTGACAAC AACGAAGTCA G

21

-76-

What is claimed is:

1. An isolated nucleic acid which encodes an acylcoenzyme A: cholesterol acyltransferase II.
5
2. An isolated nucleic acid which encodes an acylcoenzyme A: cholesterol acyltransferase III.
3. The isolated nucleic acid of claim 1 or 2, wherein
10 the nucleic acid is DNA or RNA.
4. The isolated nucleic acid of claim 3, wherein the nucleic acid is cDNA or genomic DNA.
- 15 5. The isolated nucleic acid of claim 1 comprising a nucleic acid having the sequence as set forth in Figure 15.
- 20 6. The isolated nucleic acid of claim 1, wherein the nucleic acid encodes a human wildtype acylcoenzyme A: cholesterol acyltransferase II having substantially the same amino acid sequence as set forth in Figure 15.
- 25 7. The isolated nucleic acid of claim 2, comprising a nucleic acid having the sequence as set forth in Figure 16.
- 30 8. The isolated nucleic acid of claim 2, wherein the nucleic acid encodes a human wildtype acylcoenzyme A: cholesterol acyltransferase III having substantially the same amino acid sequence as set forth in Figure 16.
- 35 9. The isolated nucleic acid of claim 1 comprising a

-77-

nucleic acid having the sequence designated Seq. I.D. No.: 11.

- 5 10. The isolated nucleic acid of claim 1, wherein the nucleic acid encodes a mouse wildtype acylcoenzyme A: cholesterol acyltransferase II having substantially the same amino acid sequence as the sequence designated Seq. I.D. No.: 12.
- 10 11. The isolated nucleic acid of claim 1, wherein the nucleic acid encodes a mutant acylcoenzyme A: cholesterol acyltransferase II.
- 15 12. The isolated nucleic acid of claim 2, wherein the nucleic acid encodes a mutant acylcoenzyme A: cholesterol acyltransferase III.
- 20 13. A vector comprising the isolated nucleic acid of claim 1 or 2.
14. The vector of claim 13 further comprising a promoter of RNA transcription operatively linked to the nucleic acid.
- 25 15. The vector of claim 14, wherein the promoter comprises a bacterial, yeast, insect or mammalian promoter.
- 30 16. The vector of claim 14, further comprising plasmid, cosmid, yeast artificial chromosome (YAC), bacteriophage or eukaryotic viral DNA.
17. The vector of claim 14 designated YEpAB-ACAT2.
- 35 18. A host vector system for the production of a

-78-

polypeptide which comprises the vector of claim 14 in a suitable host.

- 5 19. The host vector system of claim 18, wherein the suitable host is a prokaryotic or eukaryotic cell.
20. The host vector system of claim 19, wherein the prokaryotic cell is a bacterial cell.
- 10 21. The host vector system of claim 19, wherein the eukaryotic cell is a yeast, insect, plant or mammalian cell.
- 15 22. A method for producing a polypeptide which comprises growing the host vector system of claim 18 under suitable conditions permitting production of the polypeptide and recovering the polypeptide so produced.
- 20 23. A method of obtaining a polypeptide in purified form which comprises:
- (a) introducing the vector of claim 14 into a suitable host cell;
 - (b) culturing the resulting cell so as to produce the polypeptide;
 - 25 © recovering the polypeptide produced in step (b); and
 - (d) purifying the polypeptide so recovered.
- 30 24. A purified wildtype acylcoenzyme A: cholesterol acyltransferase II.
25. A purified mutant acylcoenzyme A: cholesterol acyltransferase II.

35

-79-

26. A purified wildtype acylcoenzyme A: cholesterol acyltransferase III.
- 5 27. A purified mutant acylcoenzyme A: cholesterol acyltransferase III.
- 10 28. An oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II without hybridizing to a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II.
- 15 29. An oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within the nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II without hybridizing to a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase II.
- 20 30. An oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase III without hybridizing to a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase III.
- 25 31. An oligonucleotide of at least 15 nucleotides capable of specifically hybridizing with a unique sequence of nucleotides present within the nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase III without hybridizing
- 30 35

-80-

to a nucleic acid which encodes a wildtype acylcoenzyme A: cholesterol acyltransferase III.

5 32. The oligonucleotide of claim 28, 29, 30 or 31 wherein the nucleic acid is DNA or RNA.

33. A nucleic acid having a sequence complementary to the sequence of the isolated nucleic acid of claim 1 or 2.

10

34. A method for determining whether a subject known to have an imbalance in sterol levels has the imbalance due to a defect in esterification of sterol which comprises:

15 (a) obtaining from the subject an appropriate sample containing a mixture of all of the subject's nucleic acids; and

(b) determining whether any nucleic acid in the sample from step (a) is, or is derived from, a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase so as to thereby determine whether the subject's imbalance in sterol levels is due to a defect in esterification of sterol.

25

35. The method of claim 34, wherein the determining of step (b) comprises:

(I) contacting the sample of step (a) with the isolated nucleic acid of claim 11 or 12 or the oligonucleotide of claim 29 or 31 under conditions permitting binding of any nucleic acid in the sample which is, or is derived from, a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase to the nucleic acid or

35

-81-

- oligonucleotide so as to form a complex;
- (ii) isolating the complex so formed; and
- (iii) identifying the nucleic acid in the isolated complex so as to thereby determine whether any nucleic acid in the sample contains a nucleic acid which is, or is derived from, a nucleic acid which encodes a mutant acylcoenzyme A: cholesterol acyltransferase II or III.

10

36. The method of claim 35, wherein the isolated nucleic acid or the oligonucleotide is labeled with a detectable marker.

15

37. The method of claim 36, wherein the detectable marker is a radioactive isotope, a fluorophore or an enzyme.

20

38. The method of claim 35, wherein the nucleic acid sample is first bound to a solid matrix before performing step (I).

25

39. The method of claim 35, wherein the sample comprises blood or sera.

30

40. A method for treating a subject who has an imbalance in sterol levels due to a defect in esterification of sterol which comprises introducing the isolated nucleic acid of claim 1 or 2 into the subject under conditions such that the nucleic acid expresses a wildtype acylcoenzyme A: cholesterol acyltransferase II or III, so as to thereby treat the subject.

35

41. A method for inhibiting wildtype acylcoenzyme A: cholesterol acyltransferase II or III in a subject



-82-

which comprises transforming appropriate cells from the subject with a vector which expresses the nucleic acid of claim 33, and introducing the transformed cells into the subject so as to thereby
5 inhibit wildtype acylcoenzyme A: cholesterol acyltransferase II or III.

42. The method of claim 41, wherein the nucleic acid of claim 33 is capable of specifically hybridizing to
10 a mRNA molecule encoding acylcoenzyme A: cholesterol acyltransferase II or III so as to prevent translation of the mRNA molecule.

43. A method for inhibiting the wildtype acylcoenzyme A: cholesterol acyltransferase II or III in a subject
15 which comprises introducing the oligonucleotide of claim 28 or 30 into the subject so as to thereby inhibit the wildtype acylcoenzyme A: cholesterol acyltransferase II or III.

20 44. The method of claim 43, wherein the oligonucleotide of claim 28 or 30 is capable of specifically hybridizing to a mRNA molecule encoding acylcoenzyme A: cholesterol acyltransferase II or III so as to
25 prevent translation of the mRNA molecule.

45. A method for identifying a chemical compound which is capable of inhibiting acylcoenzyme A: cholesterol acyltransferase II or III in a subject which
30 comprises:

(a) contacting a wildtype acylcoenzyme A: cholesterol acyltransferase II or III with the chemical compound under conditions permitting binding between the acylcoenzyme and the
35 chemical compound;

-83-

- (b) detecting specific binding of the chemical compound to the acylcoenzyme; and
- © determining whether the chemical compound inhibits the activity of the coenzyme so as to identify a chemical compound which is capable of inhibiting acylcoenzyme A: cholesterol acyltransferase II or III in a subject.
- 5
46. A pharmaceutical composition comprising the chemical compound identified by the method of claim 45 in an amount effective to inhibit acylcoenzyme A: cholesterol acyltransferase II or III in a subject and a pharmaceutically effective carrier.
- 10
47. A method of treating a subject who has atherosclerosis comprising administering the pharmaceutical composition of claim 46 to the subject.
- 15
48. A method of treating a subject who has hyperlipidemia comprising administering the pharmaceutical composition of claim 46 to the subject.
- 20
49. A transgenic, nonhuman mammal comprising the isolated nucleic acid of claim 1 or 2.
- 25

1/33

FIG. 1A

```

1
ARE1 MTETKDLLQD EEFLKIRRLN SAEANKRHSV TYDN..VILP QESMEVSPRS STTSLV.... EPVESTEGV ESTEAEVAG KQEQEVEYPV DAIHQKYLSH
ARE2 MDKKKOLLN EQFLRIQKLN AADAGKQSI TVDDEGELYG LDTSGNSPAN EHTATTITQN HSWVASNGDV AFIPGTATEG NTEIVTTEVI ETDDNMFKTH
hACAT MVGEEKMSLR NRLSKSRENP EEDEDQNP. .... AKESL ETPSNGRIDI KOLIAKKIKL TAEAEEL...
CON M---KDLL-- E-FLKIR-LN -ADA-KR-S- T-D-----SP-- --T----- --V-S-E-V E-----R-G KQEI--E--- -A-----H

101
ARE1 LK...SKSRS RFIHKDASKY VSFFGDVSFD PRPTLLDSAI NVPFQTTPKG PVLEKQL... ..TKTKAT VKT'VKT'KTEK
ARE2 VKT'LSSKEKA RY.KQGSBNF ISYFDDMSFE HRPSILDGVS NEPFKTKFVG PTLKEIRRR EKELMANHKN LHHRKSSPDA VDSVGKNDGA APT'VPTAAT
hACAT .....KP FFMKEVGBHF DDFVTNL.IE KSASLDNGGC AL...TTP.. SVLEGE.... KN NHRAKDLRAP PE.....Q GKIFIARRSL
CON -K---SK-K- RF-R---S-F -SFF-D-SFE -RPSLLDG-- N-PF-TTP-G PVLEKE---KN LH--K-----K-----K'IV-T---

201
ARE1 TDKADAPPGE KLESNFSGIY VFAWMFLGWI AIRCCTDYIA SYGSANWKL E IVQYMTTDLF TIAMLDLAF LCTFFVVFVH WLKKRIINW KW'GFVAVSI
ARE2 SETVVTVEIT IISNFSGLY VAFWMAIAG AVKALIDYIY QHNGSFKDSE ILKFMTTNLF TVASVDLLMY LSTYFVVGIQ YLCKWGLKW GTTGWIFTSI
hACAT LDELLEVD.. ....HIRTIV HMFIALLLIF ILSTLVWDYI DEGRVLFEFS LLSYAFGKFP TVVWTWIMF LSTFSVPYFL FQHWRTGYSK SSHPLIRSLF
CON -D---V--- ---SNFSGIY V-FWM-L--- A---L-DYI- --G-----E IL-YMTT-LF TVA--DL-MF LSTFFVW--- -L-K-----W --TG-I--SI

301
ARE1 FELAPIPVTF PIVVY..... ..YFDFNWVT RIFLFLHSV FVMKSHSPAF YNGYLWDIKQ ELEYSSKQLQ KYKESLS.PE TREILQKSCD FCLFELNYQT
ARE2 YEFLPVIFYM YLTEN..... ..ILKLHWSL KIFLFLHSLV LLMKMHSPAF YNGYLWGIKE ELQFSKSALA KYKDSINDPK VIGALEKSCE FCSFELSSQS
hACAT HGFLPMIFQI GVLGFGPTYV VLAYTLPPAS RFIIFEQIR FVMKAHSP.. ....VRE NV...PRVLN SAKESSTVP IPTVNO.... ..
CON -EFLP-IF-- ----- -L-W-S RIFLFLHS-V FVMK-HSPAF YNGYLW-IKE EL--S---L- KYKES-S-P- ----LQKSC- FC-FEL--Q-

```

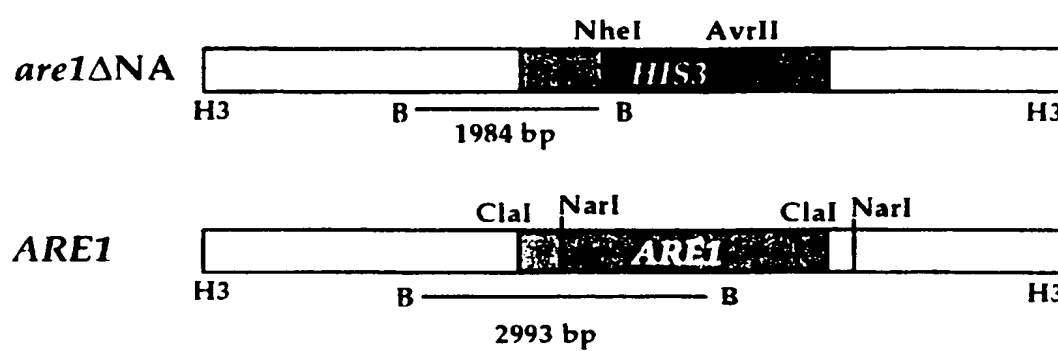
2/33

601

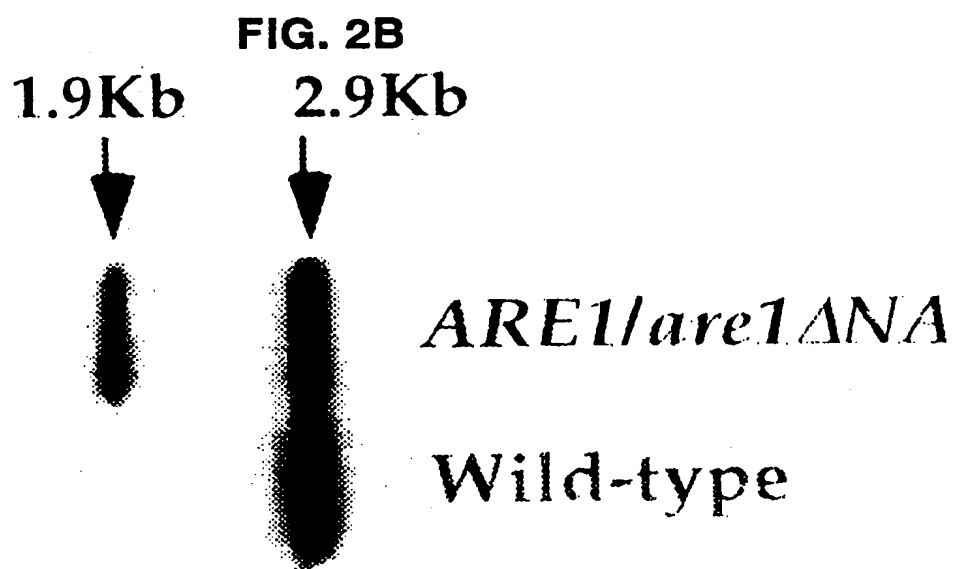
ARE1	ARE2	HACAT	CON
VRGYLPMFQL SQF VMTALS	NRGYLPMFQL SQF VMTALS	NRGYLPMFQL SQF VMTALS	NRGYLPMFQL SQF VMTALS
LRFLYLPFFQMQ LQ MPLVALT	LRFLYLPFFQMQ LQ MPLVALT	LRFLYLPFFQMQ LQ MPLVALT	LRFLYLPFFQMQ LQ MPLVALT
FYPVLPLVLFM FFGMAFNPIV	FYPVLPLVLFM FFGMAFNPIV	FYPVLPLVLFM FFGMAFNPIV	FYPVLPLVLFM FFGMAFNPIV
-R-YLP-FQ- -Q-M- -AL-	-R-YLP-FQ- -Q-M- -AL-	-R-YLP-FQ- -Q-M- -AL-	-R-YLP-FQ- -Q-M- -AL-
NTKF-R-R- NTKF-R-R-	NTKF-R-R- NTKF-R-R-	NTKF-R-R- NTKF-R-R-	NTKF-R-R- NTKF-R-R-
I-NV-FW-G- C-GPSV--TL	I-NV-FW-G- C-GPSV--TL	I-NV-FW-G- C-GPSV--TL	I-NV-FW-G- C-GPSV--TL
FLNGVLLLCF YSQEWYARRHC	FLNGVLLLCF YSQEWYARRHC	FLNGVLLLCF YSQEWYARRHC	FLNGVLLLCF YSQEWYARRHC
IWNVLMTSL	IWNVLMTSL	IWNVLMTSL	IWNVLMTSL
CMGSPVMCTL YLTF2	CMGSPVMCTL YLTF2	CMGSPVMCTL YLTF2	CMGSPVMCTL YLTF2
IGNVIFWLGI	IGNVIFWLGI	IGNVIFWLGI	IGNVIFWLGI
LSNVVFSFGV	LSNVVFSFGV	LSNVVFSFGV	LSNVVFSFGV
CSGSIIMTL YLTL	CSGSIIMTL YLTL	CSGSIIMTL YLTL	CSGSIIMTL YLTL
NTKFLRARPQ	NTKFLRARPQ	NTKFLRARPQ	NTKFLRARPQ
YVPRSWTCRY VF	YVPRSWTCRY VF	YVPRSWTCRY VF	YVPRSWTCRY VF

3/33

FIG. 2A



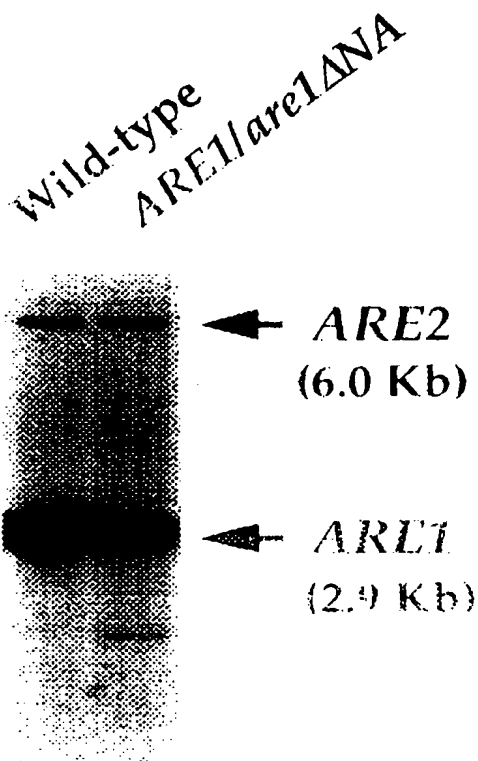
4/33



SUBSTITUTE SHEET (RULE 26)

5/33

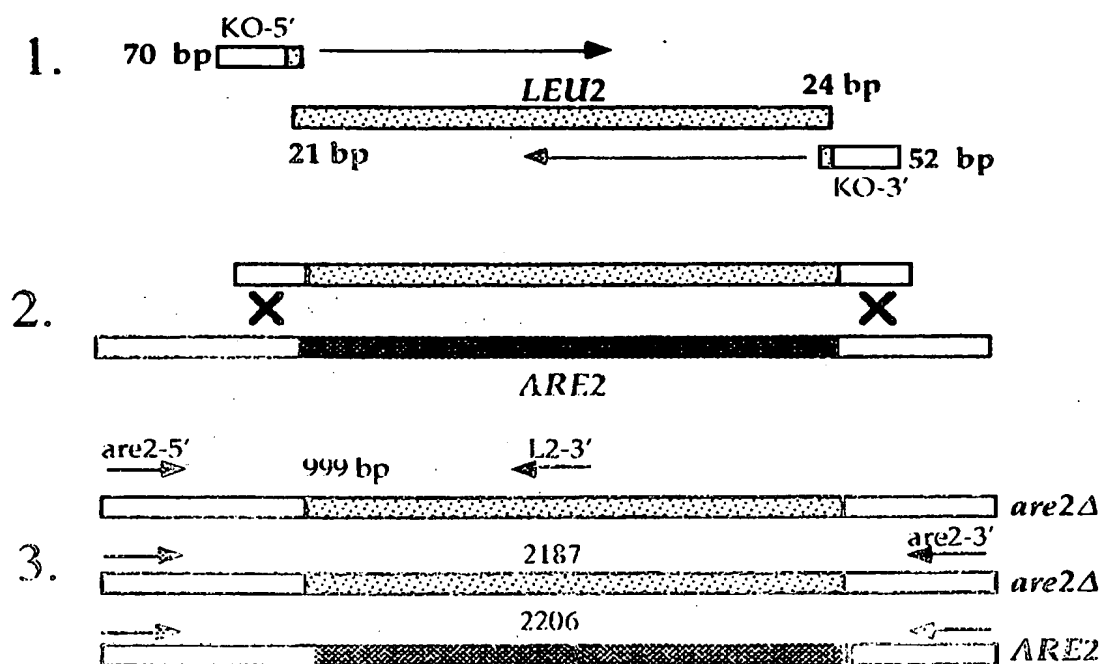
FIG. 2C



SUBSTITUTE SHEET (RULE 26)

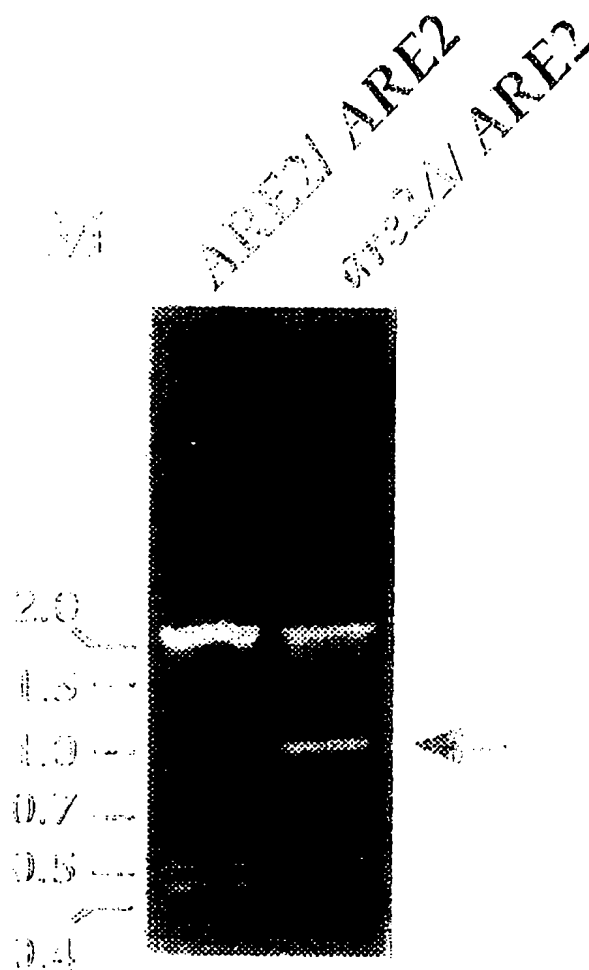
6/33

FIG. 2D



7/33

FIG. 2E



8/33

FIG. 3A
Wild-type

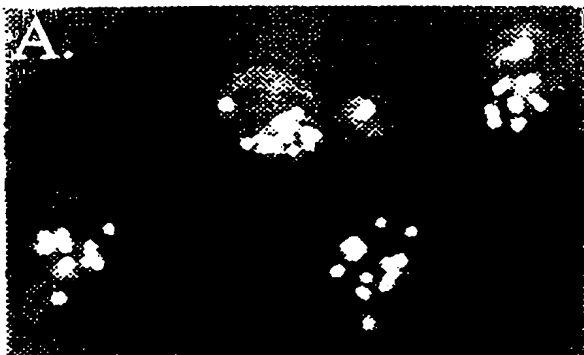
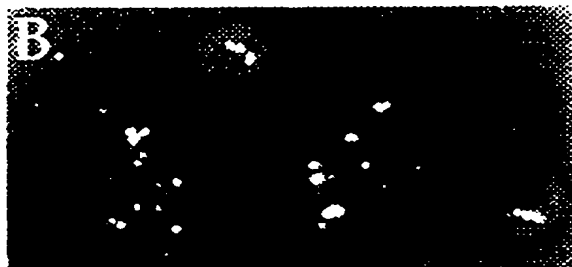
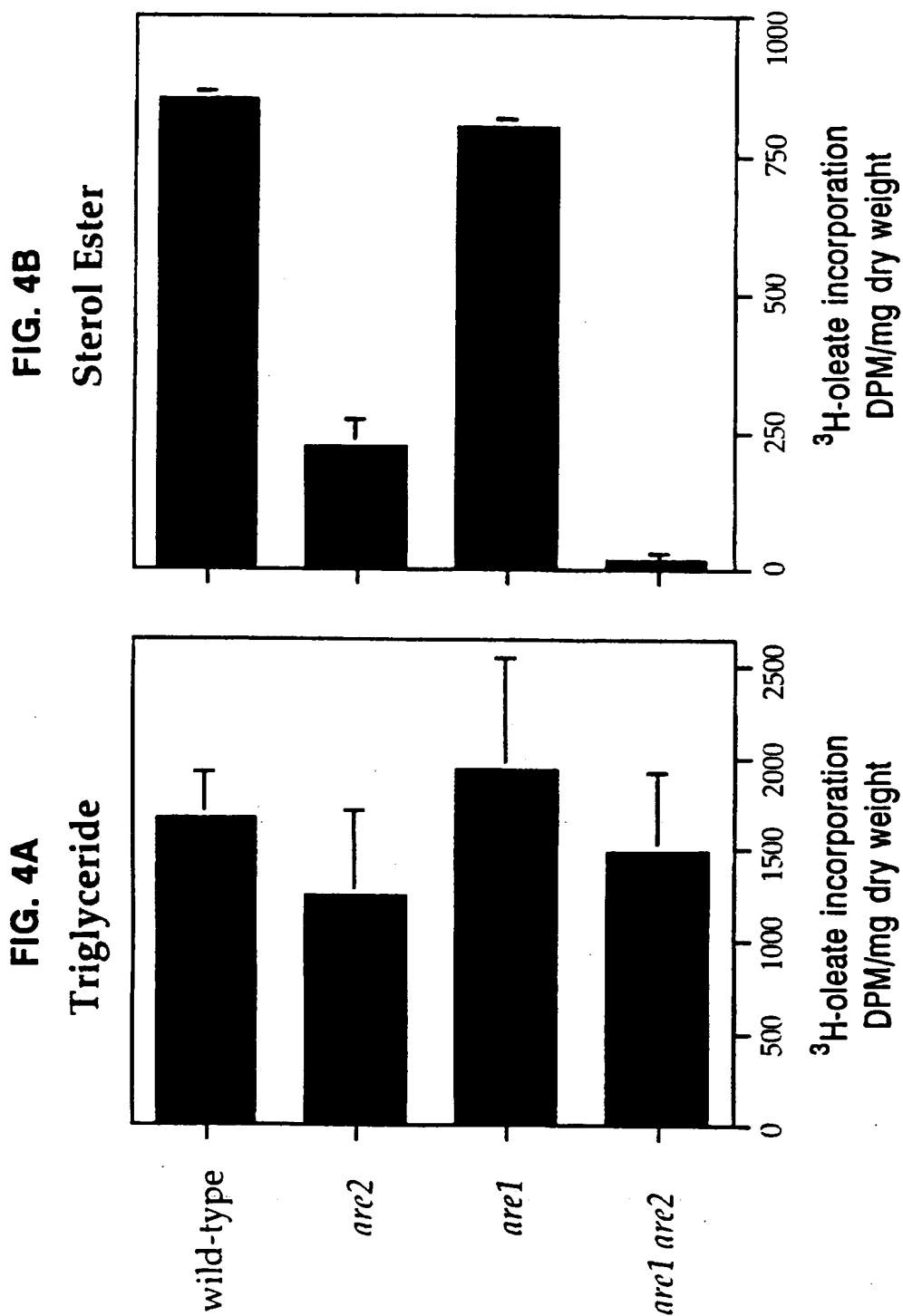


FIG. 3B
are1 are2 mutant

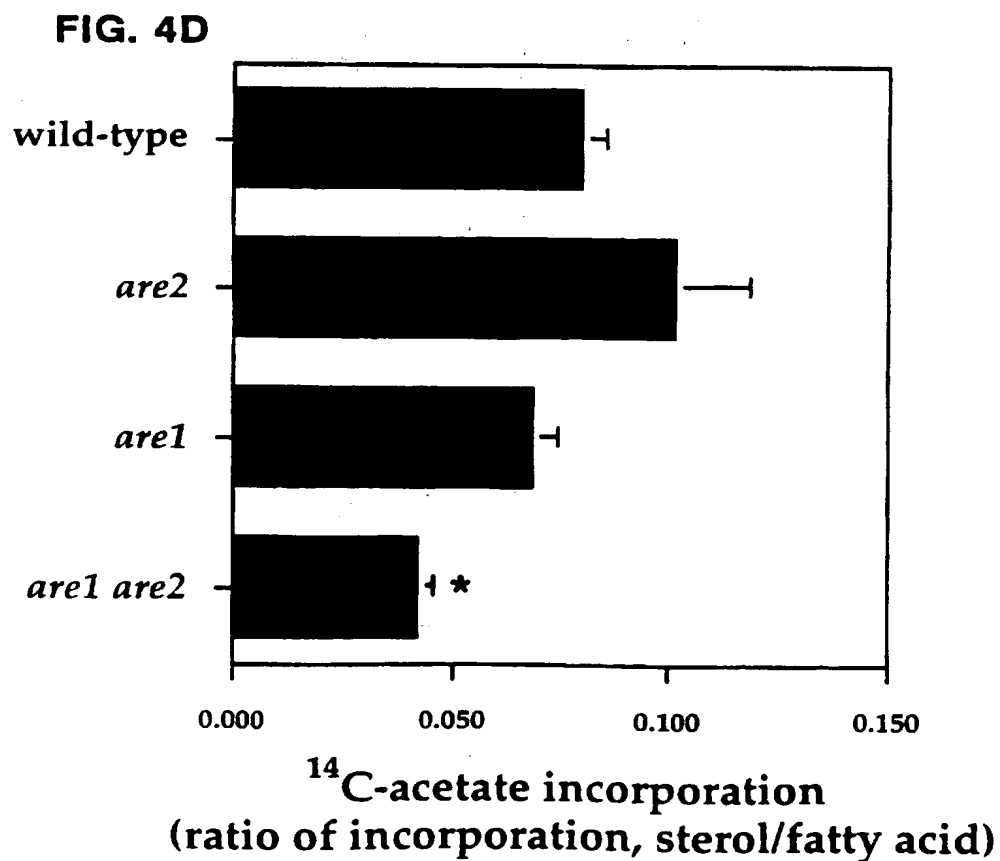
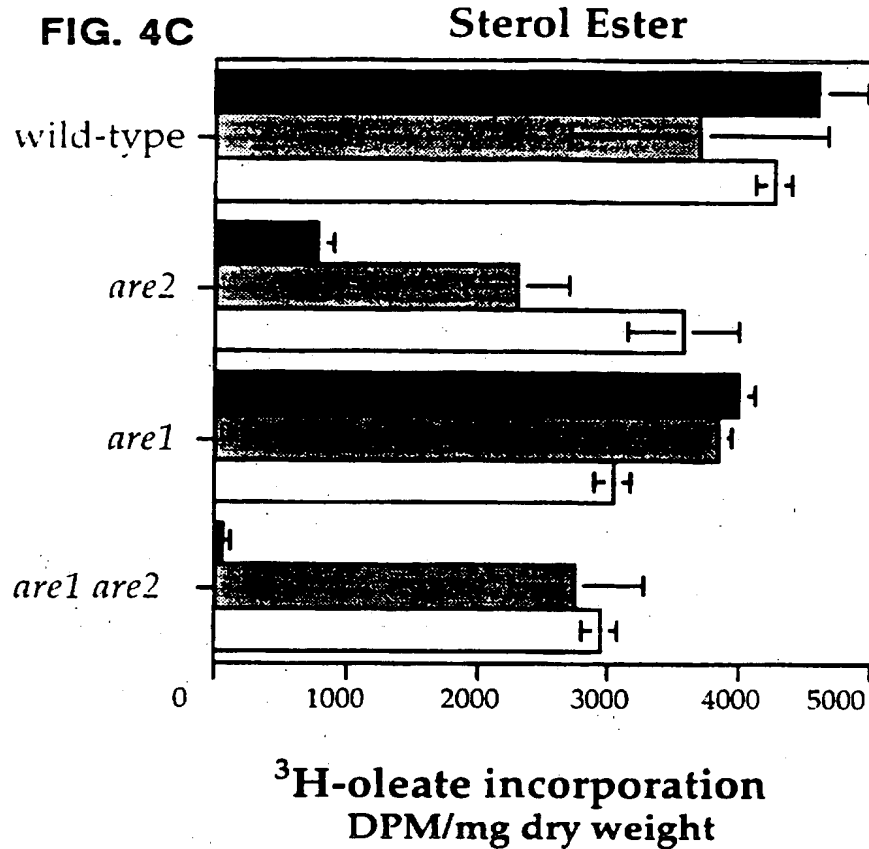


SUBSTITUTE SHEET (RULE 26)

9/33



10/33



11/33

FIG. 5A-1

```

1  gggtagagacggggttcaccgtgttagccaggatggtctggatctcctgacctcgtgatccaccaccctcgccctcctaaagtctgggattacagaca 100
101  tgagccaccgcccagccctattcatccctttcaaaagtcagaccctaggaagctggaggaggtgggcatggttttacagtgaatttctgatttca 200
201  ctcagggtgataaatcagactcttggggaagcgggtggtggctctggacagcagcaggaatggggatccagttagcaacaaatccatggacctatgacag 300
301  gctgaaagccaccctctccatcttgggaggttgccaatgtctgatttaacactatccaatgaatgatcattgaaagtaaaaaaataaactatcaactag 400
401  cagaaaaatataaatggtaagcattagcacatatttcacatgttttatatttggctctcagattgacctataaaacaaagtctgggaaattctatatgatcc 500
501  tgaaaaaatgatacgtcgtctggatggtagaataaagttggagaaatgtttaagccaaaatgcagctcttaccatgactttttattttattttaatt 600
601  ttcaggatttttggtatcacagggtgttttggttacatggaaaaagttcttactggtgatttctgagatttttagttcacccctatcctgagcagtgta 700
701  actgttcccaatatgtagccttttatccctcacccctcctaagttcaagaagactatggtcctgcagaaaagcttttatatgtaattaacatatctttatct 800
801  ttatctttataggcagtagactcatcttttgaacagattccattaaagtgatgaatgtgtacctccctctagcctttattattactgtttttgctattac 900
901  atgtgttagtgtgaatttaaatgttaaaatgtatcccatgtgctactatggcaaaaggttgactcataagagtttagcacgggttaagatctgaa 1000
1001  agttttctccagcctcttatcactggcgagacttcacaattcatggaagccaccagtgagatgacattgcctcaggcagttactatttttatattcta 1100
1101  taactcgaggagtcagggttttcggaatcattaaacttttttgtcctttttaaagttggagacagcaattgtagacagccttcagtggttatctttt 1200
1201  tgtgtctccttacctgtggagaagcctattagctgggatatgtagttaaatagctatatatttatatatatccagggcaccccggaattcgggagagcttccc 1300
1301  ggagtcgaccttctctgctgtgaccttcccggtctgacctcttggccgaagtgcccgctgcggggcgcggtcctcagacaataca  ATG 1399
1  1
1400  GTG GAT GAG AAG ATG TCT CTA AGA AAC CGG CTG TCA AAG TCC AGG GAA AAT CCT GAG GAA GAT GAA GAC CAG 1474
2  V G E K M S L R N R L S K S R E N P E E D Q 26
1475  ACA AAC CCT GCA AAG GAG TCC CTA GAG ACA CCT AGT AAT GGT CGA ATT GAC ATA AAA CAG TTG ATA GCA AAG AAG 1549
27  R N P A K E S L E T P S N G R I D I K Q L I A K K 51
1550  ATA AAG TTG ACA GCA GAG GAA TTG AAG CCA TTT TTT ATG AAG GAA GTT GGC AGT CAC TTT GAT GAT TTT 1624
52  I K L T A E A E L K P F M K E V G S H F D D F 76

```


12/33

FIG. 5A-2

1625 GTG ACC AAT CTC ATT GAA AAG TCA GCA TCA TTA GAT AAT GGT GGG TGC GCT CTC ACA ACC TTT TCT GTT CTT GAA 1699
 77 V T N L I E K S A S L D N G G C A L T T F S V L E 101
 1700 GGA GAG AAA AAC CAT AGA GCG AAG GAT TTG AGA GCA CCT CCA GAA CAA GGA AAG ATT TTT ATT GCA AGG CGC 1774
 102 G E K N N H R A K D L R A P P E Q G K I F I A R R 126
 1775 TCT CTC TTA GAT GAA CTG CTT GAA GTG GAC CAC ATC AGA ACA ATA TAT CAC ATG TTT ATT GCC CTC CTC ATT CTC 1849
 127 S L L D E L L E V D H I R T I Y H M F I A L L I L 151
 1850 TTT ATC CTC AGC ACA CTT GTA GAT TAC TAC GAT GAA GGA AGG CTG GTG CTT GAG TTC AGC CTC CTG TCT TAT 1924
 152 F I L S T L V V D Y I D E G R L V L E F S L L S Y 176
 1925 GCT TTT GGC AAA TTT CCT ACC GTT GTT TGG ACC TGG TGG ATC ATG TTC CTG TCT ACA TTT TCA GTT CCC TAT TTT 1999
 177 A F G K F P T V V T W T W W I M F L S T F S V P Y F 201
 2000 CTG TTT CAA CAT TGG CGC ACT GGC TAT AGC AAG AGT TCT CAT CCG CTG ATC CGT TCT CTC TTC CAT GGC TTT CTT 2074
 202 L F Q H W R T G Y S K S S H P L I R S L F H G F L 226
 2075 TTC ATG ATC TTC CAG ATT GGA GTT CTA GGT TTT GGA CCA ACA TAT GTT GTG TTA GCA TAT ACA CTG CCA CCA GCT 2149
 227 F M I F Q I G V L G F G P T Y V V L A Y T L P P A 251
 2150 TCC CGG TTC ATC ATT ATA TTC GAG CAG ATT CGT TTT GTA ATG AAG GCC CAC TCA TTT GTC AGA GAG AAC GTG CCT 2224
 252 S R F I I I F E Q I R F V M K A H S F V R E N V P 276
 2225 CGG GTA CTA AAT TCA GCT AAG GAG AAA TCA AGC ACT GTT CCA ATA CCT ACA GTC AAC CAG TAT TTG TAC TTC TTA 2299
 277 R V L N S A K E K S S T V P I P T V N Q Y L Y F L 301
 2300 TTT GCT CCT ACC CTT ATC TAC CGT GAC AGC TAT CCC AGG AAT CCC ACT GTA AGA TGG GGT TAT GTC GCT ATG AAG 2374
 302 F A P T L I Y R D S Y P R N P T V R W G Y V A M K 326
 2375 TTT GCA CAG GTC TTT GGT TGC TTT TTC TAT GTG TAC TAC ATC TTT GAA AGG CTT TGT GCC CCC TTG TTT CGG AAT 2449
 327 F A Q V F G C F F Y V Y Y I F E R L C A P L F R N 351
 2450 ATC AAA CAG GAG CCC TTC AGC GCT CGT GTT CTG GTC CTA TGT GTA TTT AAC TCC ATC TTG CCA GGT GTG CTG ATT 2524
 352 I K Q E P F S A R V L V L C V F N S I L P G V L I 376

13/33

FIG. 5A-3

2525 CTC TTC CTT ACT TTT TTT GCC TTT TTG CAC TGC TGG CTC AAT GCC TTT GCT GAG ATG TTA CGC TTT GGT GAC AGG 2599
 377 L F L T F A F L H C W L N A F A E M L R F G D R 401

2600 ATG TTC TAT AAG GAT TGG TGG AAC TCC ACG TCA TAC TCC AAC TAT TAT AGA ACC TGG AAT GTG GTG GTC CAT GAC 2674
 402 M F Y K D W N S T S Y S N Y Y R T W N V V V H D 426

2675 TGG CTA TAT TAC TAT GCT TAC AAG GAC TTT CTC TGG TTT TTC TCC AAG AGA TTC AAA TCT GCT GCC ATG TTA GCT 2749
 427 W L Y Y A Y K D F L W F F S K R F K S A A M L A 451

2750 GTC TTT GCT GTA TCT GCT GTA CAC GAA TAT GCC TTG GCT GTT TGC TTG AGC TTT TTC TAT CCC GTG CTG TTC 2824
 452 V F A V S A V V H E Y A L A V C L S F F Y P V L F 476

2825 GTG CTC TTC ATG TTC TTT GGA ATG GCT TTC AAC TTC ATT GTC AAT GAT AGT CCG AAA AAG CCG ATT TGG AAT GTT 2899
 477 V L F M F F G M A F N F I V N D S R K K P I W N V 501

2900 CTG ATG TGG ACT TCT CTT TTC TTG GGC AAT GGA GTC TTA CTC TGC TTT TAT TCT CAA GAA TGG TAT GCA CGT CGG 2974
 502 L M W T S L F L G N G V L L C F Y S Q E W Y A R R 526

2975 CAC TGT CCT CTG AAA AAT CCC ACA TTT TTG GAT TAT GTC CGG CCA CGT TCC TGG ACT TGT CGT TAC GTC TTT TAG 3049
 527 H C P L K N P T F L D Y V R P R S W T C R Y V F * 551

3050 aagcttggaactttgttccctctgactgaagattgggttagctccctgatttggagccagctgtttccagttgttactgaagttatctgtgtatttg 3149

3150 gaccactccaggctttacagatgactcactccattccttaggtcacttgaagccaaactgttgaagttcactggagttctgtacacttaagcagagcaga 3249

3250 actttttttgtgggctgggtggggggaagaccgactaacagctgaagtaatgacagattgttgcgtgggtcataatcagctttatcccttgtaattat 3349

3350 atctgtttttgttcttgactctgtccaatcagagaataaacatcatagtttcttggccactgaattagccaaaacacttaggaagaatacacttaaatatc 3449

3450 ctctggcttagaaaattttttcatgcacactgttggaaatgtatgctaatgaacatgcaattggggaagaaaaaatgagaatgattttgtctattcttag 3549

3550 tagaaagaaaatgtctgttttccaaaagataatgtttatacatcctatttttgaattttttgaaaaaaagtttcaatgttccagttttcccttagttttacctt 3649

FIG. 5B-1

1 gccctccagctcttactaagaccggtcggaagcatgctggcgcatatataccaaaccacacacacatggctcctcctcgctgcataaaatctccccagac 200

101 agtcgggaccgccaccccgatataccagaatgaaactgcacggctgcagattcaaaagctccaacggcctcagcgtcatcttcgctcggtatgctgcactc 200

201 tggctgaaccgctgtactgtgtgcttgcctatcattatagaaaatctccggtggcgccaaactcctcaggacgtgacattatttcttctctgatattt 300

301 cctgtgtttccgtaaccgcaccttttttagcactacttttttactatgctcttcttctctgcttcttctctgttttttcttcttccactatgtatgtg 400

401 ctgctcatctcttcttttatcgataaaaattgaaaatgtgagatgggtgtagagtgaiaaaaaaataatctggcttggccatcaaataccggccgt 500

501 ggltggactcgttttagcgaacaatagcaccacccagcagccctggcaacatgcggatgatataagaaggacgagcgtggtgaggaaaagggcgccattggc 600

601 acactcacgcaggtggttcttcagcacggttgcagcaagcgttcgcaaaacagattgcaaga ATG ACG GAG ACT AAG GAT TTG TTG CAA 689
1 M T E T K D L L Q 9

690 GAC GAA GAG TTT CTT AAG ATC CGC AGA CTC AAT TCC GCA GAA GCC AAC AAA CGG CAT TCG GTC ACG TAC GAT AAC 764
10 D E E F L K I R R L N S A E A N K R H S V T Y D N 34

765 GTG ATC CTG CCA CAG GAG TCC ATG GAG TCC ATG GAG GTT TCG CCA CGG TCG TCT ACC ACG TCG CTG GTG GAG CCA GTG GAG TCG 839
35 V I L P Q E S M E V S P R S S T T S L V E P V E S 59

840 ACT GAA CGA GTG GAG TCG ACT GAG GCG GAA CGT GTG GCA GGG AAG CAG GAG CAG GAG GAG TAC OCT GTG GAC 914
60 T E G V E S T E A E R V A G K Q E E E Y P V D 84

915 GCC CAC ATG CAA AAG TAC CTT TCA CAC CTG AAG AGC AAG TCT CGG TCG AGG TTC CAC CGA AAG GAT GCT AGC AAG 989
85 A H M Q K Y L S H L K S K S R S R F H R K D A S K 109

990 TAT GTG TCG TTT TTT GGG GAC GTG AGT TTT GAT CCT CGC CCC ACG CTC CTG GAC AGC GCC ATC AAC GTG CCC TTC 1064
110 Y V S F F G D V S F D P R P T L L D S A I N V P F 134

1065 CAG ACG ACT TTC AAA GGT CCG GTG CTG GAG AAA CAG CTC AAA AAT TTA CAG TTG ACA AAG ACC AAG ACC AAG GCC 1139
135 Q T T F K G P V L E K Q L K N L Q L T K T K T K A 159

1140 ACG GTG AAG ACT ACG GTG AAG ACT ACG GAG AAA ACG GAC AAG GCA GAT GCC CCC CCA GGA GAA AAA CTG GAG TCG 1214
160 T V K T T V K T T E K T D K A D A P P G E K L E S 184

1215 AAC TTT TCA GGG ATC TAC GTG TTC GCA TGG ATG TTC TTG GGC TGG ATA GCC ATC AGG TGC TGC ACA GAT TAC TAT 1289
185 N F S G I Y V F A W M F L G W I A I R C C T D Y Y 209

15/33

FIG. 5B-2

1290 GCG TCG TAC GGC AGT GCA TGG AAT AAG CTG GAA ATC GTG CAG TAC ATG ACA ACG GAC TTG TTC ACG ATC GCA ATG 1364
 210 A S Y G S A W N K L E I V Q Y M T T D L F T I A M 234
 1365 TTG GAC TTG GCA ATG TTC CTG TGC ACT TTC TTG GTG GTT TTC GTG CAC TGG CTG GTG AAA AAG CGG ATC ATC AAC 1439
 235 L D L A M F L C T F F V V F V H W L V K K R I I N 259
 1440 TGG AAG TGG ACT GCG TTC GTT GCA GTG AGC ATC TTC GAG TTG GCT TTC ATC CCC GTG ACG TTC CCC ATT TAC GTC 1514
 260 W K W T G F V A V S I F E L A F I P V T F P I Y V 284
 1515 TAC TAC TTT GAT TTC AAC TGG GTC ACG AGA ATC TTC CTG TTC CTG CAC TCC GTG GTG TTT GTT ATG AAG AGC CAC 1589
 285 Y Y F D F N W V T R I F L F L H S V V F V M K S H 309
 1590 TCG TTT GCC TTT TAC AAC GGG TAT CTT TGG GAC ATA AAG CAG GAA CTC GAG TAC TCT TCC AAA CAG TTG CAA AAA 1664
 310 S F A F Y N G Y L W D I K Q E L E Y S S K Q L Q K 334
 1665 TAC AAG GAA TCT TTG TCC CCA GAG ACC CGC GAG ATT CTG CAA AAA AGT TGC GAC TTT TGC CTT TTC GAA TTG AAC 1739
 335 Y K E S L S P E T R E I L Q K S C D F C L F E L N 359
 1740 TAC CAG ACC AAG GAT AAC GAC TTC CCC AAC AAT TGC AGC AAT TTC ATG TTC TGT TGT TTC CCC GTC 1814
 360 Y Q T K D N D F P N N I S C S N F F M F C L F P V 384
 1815 CTC GTG TAC CAG ATC AAC TAC CCA AGA ACG TCG CGC ATC AGA TGG AGG TAT GTG TTG GAG AAG GTG TGC GCC ATC 1889
 385 L V Y Q I N Y P R T S R I R W R Y V L E K V C A I 409
 1890 ATT GGC ACC ATC TTC CTC ATG ATG GTC ACG GCA CAG TTC TTG CAC CCG GTG GCC ATG CGC TGT ATC CAG TTC 1964
 410 I G T I F L M M V T A Q F F M H P V A M R C I Q F 434
 1965 CAC AAC ACG ACC TTC GGC GGC TGG ATC CCC GCG ACG CAA GAG TGG TTC CAC CTG CTC TTC GAC ATG ATT CCG 2039
 435 H N T P T F G G G W I P A T Q E W F H L L F D M I P 459
 2040 GGC TTC ACT GTT CTG TAC ATG CTC ACG TTT TAC ATG ATA TGG GAC CCT TTA TTG AAT TGC GTG GCG GAG TTG ACC 2114
 460 G F T V L Y M L T F Y M I W D A L L N C V A E L T 484

16/33

FIG. 5B-3

```

2115 AGG TTT GCG GAC AGA TAT TTC TAC TAC GGC GAC TGG TGG AAT TGC GTT TCG TTT GAA GAG TTT AGC AGA ATC TGG AAC 2189
485 R F A D R Y F Y G D W W N C V S F E E F S R I W N 509

2190 GTC CCC GTT CAC AAA TTT TTA CTA AGA CAC GTG TAC CAC AGC TCC ATG GGC GCA TTG CAT TTG AGC AAG AGC CAA 2264
510 V P V H K F L L R H V Y H S S M G A L H L S K S Q 534

2265 GCT ACA TTA TTT ACT TTT TTC TTG AGT GCC GTG TTC CAC GAA ATG GCC ATG TTC GCC ATT TTC AGA AGG GTT AGA 2339
535 A T L F T F F L S A V F H E M A M F A I F R R V R 559

2340 GGA TAT CTG TTC ATG TTC CAA CTG TCG CAG TTT GTG TGG ACT GCT TTG AGC AAC ACC AAG TTT CTA CCG GCA AGA 2414
560 G Y L F M F Q L S Q F V W T A L S N T K F L R A R 584

2415 CCG CAG TTG TCC AAC GTT GTC TTT TCG TTT GGT GTC TGT TCA GGG CCC AGT ATC ATT ATG ACG TTG TAC CTG ACC 2489
585 P Q L S N V V F S F G V C S G P S I I M T L Y L T 609

2490 TTA TGA actgccaccataccacgtgtgtccctcgcaagcccttgatagatatacaatagggaatgggcgtccgtccaccgtggtcacaagacaggggca 2587
610 L * 611

2588 aagagctcctaggt 2601

```

17/33

FIG. 5C-1

1 tataaaattccttcatcaatacatctatatattcgaatatatagataaaccatacaaaaacatactgaaatttttggaaaaaacactaaaaactattcat 100
 101 tgcagttacacgtgaatgctaaactttatatcgtctctgtcgtcccgaggagttacatttaacggtcttcgcgcaataaccggaataattccaacag 200
 201 tttcttttgaataattattaagccttctttttcccggaatctataagaggggacgaaaattagccgctatttaattctgtgtattgccacctagacaagaag 300
 301 taaacagacacattacgttagcaaaagcaacaatacaacaacaacc ATG GAC AAG AAG AAG GAT CTA CTG GAG AAC GAA CAA TTT 386
 1 M D K K K D L L E N E Q F 13
 387 CTC CGC ATC CAA AAG CTC AAC GCT GCC GAT GCG GGC AAA AGA CAA TCT ATA ACA GTG GAC GAC GAG GGC GAA CTA 461
 14 L R I Q K L N A A D A G K R Q S I T V D D E G E L 38
 462 TAT GGG TTA GAC ACC TCC GGC AAC TCA CCA GCC AAT GAA CAC ACA GCT ACC ACA ATT ACA CAG AAT CAC AGC GTG 536
 39 Y G L D T S G N S P A N E H T A T T I T Q N H S V 63
 537 GTG GCC TCA AAC GGA GAC GTC GCA TTC ATC CCA GGA ACT GCT ACC GAA GGC AAT ACA GAG ATT GTA ACT GAA GAA 611
 64 V A S N G D V A F I P G T A T E G N T E I V T E E 88
 612 GTG ATT GAG ACC GAT AAC ATG TTC AAG ACC CAT GTG AAG ACT TTA AGC TCC AAA GAG AAG GCA CGG TAT AGG 686
 89 V I E T D D N M F K T H V K T L S S K E K A R Y R 113
 687 CAA GGG TCC TCT AAC TTT ATA TCG TAT TTC GAT GAT ATG TCA TTT GAA CAC AGG CCC AGT ATA TTA GAT GCG TCA 761
 114 Q G S S N F I S Y F D M S F E H R P S I L D G S 138
 762 GTT AAC GAG CCC TTC AAG ACC AAA TTC GTG GGA CCT ACT TTA GAA AAG GAG ATC AGA AGG GAG AAA GAG CTA 836
 139 V N E P F K T K F V G P T L E K E I R R E K E L 163
 837 ATG GCC ATG CGC AAA AAT TTA CAC CAC CGC AAG TCC TCC CCA GAT GCT GTC GAC TCA GTA GGG AAA AAT GAT GGC 911
 164 M A M R K N L H H R K S S P D A V D S V G K N D G 188
 912 GCC GCC CCA ACT ACT GTT CCA ACT GCC ACC TCA GAA ACG GTG GTC ACC GTT GAA ACC ACC ATA ATT TCA TCC 986
 189 A A P T T V P T A A T S E T V V T V E T T I I S S 213
 987 AAT TTC TCC GGG TTG TAC GTG GCG TTT TGG ATG GCT ATT GCA TTT GGT GCT GTC AAG GCT TTA ATA GAC TAT TAT 1061
 214 N F S G L Y V A F W M A I A F G A V K A L I D Y Y 238

18/33

FIG. 5C-2

1062 TAC CAG CAT AAT GGT AGC TTC AAG GAT TCG GAG ATC TTG AAA TTT ATG ACT ACG AAT TTG TTC ACT GTG CCA TCC 1136
 239 Y Q H N G S F K D S E I L K F M T T N L F T V A S 263
 1137 GTA GAT CTT TTG ATG TAT TTG AGC ACT TAT TTT GTC GTT GGA ATA CAA TAC TTA TGC AAG TCG GGG GTC TTG AAA 1211
 264 V D L L M Y L S T Y F V G I Q Y L C K W G V L K 288
 1212 TGG GGC ACT ACC GGC TGG ATC TTC ACC TCA ATT TAC GAG TTT TTG TTT GTT ATC TTC TAC ATG TAT TTA ACA GAA 1286
 289 W G T T G W I F T S I Y E F L F V I F Y M Y L T E 313
 1287 AAC ATC CTA AAA CAC TGG CTG TCC AAG ATC TTC CTT TTT TTG CAT TCT TTA GTT TTA TTG ATG AAA ATG CAT 1361
 314 N I L K L H W L S K I F L F L H S L V L L M K M H 338
 1362 TCT TTC GCC TTC TAC AAT GGC TAT CTA TGG GGT ATA AAG GAA GAA CTA CAA TTT TCC AAA AGC GCT CTT GCC AAA 1436
 339 S F A F Y N G Y L W G I K E E L Q F S K S A L A K 363
 1437 TAC AAG GAT TCT ATA AAT GAT CCA AAA GTT ATT GGT GCT CTT GAG AAA AGC TGT GAG TTT TGT AGT TTT GAA TTG 1511
 364 Y K D S I N D P K V I G A L E K S C E F C S F E L 388
 1512 AGC TCT CAG TCT TTA AGC GAC CAA ACT CAA AAA TTC CCC AAC AAT ATC AGT GCA AAA AGC TTT TTT TGG TTC ACC 1586
 389 S S Q S L S D Q T Q K F P N N I S A K S F F W F T 413
 1587 ATG TTT CCA ACC CTA ATT TAC CAA ATT GAA TAT CCA AGA ACT AAG GAA ATC AGA TGG AGC TAC GTA TTA GAA AAG 1661
 414 M F P T L I Y Q I E Y P R T K E I R W S Y V L E K 438
 1662 ATC TGC GCC ATC TTC GGT ACC ATT TTC TTA ATG ATG ATA GAT GCT CAA ATC TTG ATG TAT CCT GTA GCA ATG AGA 1736
 439 I C A I F G T I F L M M I D A Q I L M Y P V A M R 463
 1737 GCA TTG GCT GTG CGC AAT TCT GAA TGG ACT GGT ATA TTG GAT AGA TTA TTG AAA TGG GTT GGA TTG CTC GTT GAT 1811
 464 A L A V R N S E W T G I L D R L L K W V G L L V D 488
 1812 ATC GTC CCA GGG TTT ATC GTG TTT ATC TTG GAC TTC TAT TTG ATT TGG GAT GCC ATT TTG AAC TGT GTG GCT 1886
 489 I V P G F I V M Y I L D F Y L I W D A I L N C V A 513
 1887 GAA TTG ACA AGA TTT GGC GAC AGA TAT TTC TAC GGT GAC TGG TGG AAT TGT GTT AGT TCG GCA GAC TTC AGT AGA 1961
 514 E L T R F G D R Y F Y G D W W N C V S W A D F S R 538

19/33

FIG. 5C-3

```

1962 ATT TGG AAC ATC CCA GTG CAT AAG TTT TTG TTA AGA CAT GTT TAC CAT AGT TCA ATG AGT TCA TTC AAA TTG AAC 2036
539 I W N I P V H K F L L R H V Y H S S M S S F K L N 563

2037 AAG AGT CAA GCA ACT TTG ATG ACC TTT TTC TTA AGT TCC GTC GTT CAT GAA TTA GCA ATG TAC GTT ATC TTC AAG 2111
564 K S Q A T L M T F F L S S V V H E L A M Y V I F K 588

2112 AAA TTG AGG TTT TAC TTG TTC TTC CAA ATG CTG CAA ATG CCA TTA GTA GCT TTA ACA AAT ACT AAA TTC ATG 2186
589 K L R F Y L F F Q M L Q M P L V A L T N T K F M 613

2187 AGG AAC AGA ACC ATA ATC GGA AAT GTT ATT TTC TGG CTC GGT ATC TGC ATG GGA CCA AGT GTC ATG TGT ACG TTG 2261
614 R N R T I I G N V I F W L G I C M G P S V M C T L 638

2262 TAC TTG ACA TTC TAA ggcacccctgcaactgttctgtggagctattaaatctttatagtaaattttttttacttttttttttttttttttttt 2356
639 Y L T F * 643

2357 ttattattacaagcgtctatatattttctattatagaatatattgtcatttattacattggttca 2421

```


20/33

FIG. 5D

```

1  ATG GAC AAC GCG GCG TCT GAT ACG ACT CAC TAT AGG GAA TTT GGC CCT CGA GCA GTA GAT TCG GCA CGA TGG GCA 75
1  M  D  N  A  G  S  D  T  T  H  Y  R  E  F  G  P  R  A  V  D  S  A  R  W  A  25

76 CGA GGA CTC CAT CAT GTT CCT CAA GGT TTA TTC CTA CCG GGA TGT CAA CCT GTG GTG CCG CCA GCG AAG GGT CAA 150
26 R  G  L  H  H  V  P  Q  A  L  F  L  P  G  C  Q  P  V  V  P  P  A  K  G  Q  50

151 GGC CAA AGC TGT CTC TAC AGG GAA GAA GGT CAG TGG GGC TGC TGC GAG CAA GCT GTG AGC TAT CCA GAC AAC CTG 225
51 G  Q  S  C  L  Y  R  E  E  G  Q  W  G  C  C  E  Q  A  V  S  Y  P  D  N  L  75

226 ACC TAC CGA GAT CTC GAT TAC TTC ATC TTT GCT CCT ACT TTG TGT TAT GAA CTC AAC TTT CCT CCG TCC CCC CGA 300
76 T  Y  R  D  L  D  Y  F  I  F  A  P  T  L  C  Y  E  L  N  F  P  R  S  P  R  100

301 ATA CGA GAG CGC TTT CTG CTA CGA GAT CTT GAG ATG CTC TTT TTT ACC CAG CTT CAA GTG GCG CTG ATC CAA 375
101 I  R  E  R  F  L  L  L  R  V  L  E  M  L  F  F  T  Q  L  Q  V  G  L  I  Q  125

376 CAG TGG ATG GTC CCT ACT ATC CAG AAC TCC ATG GAA GCC CTT TCA AGA GCT TCT GGC AGT TTT GGA GAC CGC GAG 450
126 Q  W  M  V  P  T  I  Q  N  S  M  E  A  L  S  R  A  S  G  S  F  G  D  R  E  150

451 TTC TAC AGA GAT TGG TGG AAT GCT GAG TCT GTC ACC GAC TTT TGG CAG AAC TGG AAT ATC CCC GTG G 517
151 F  Y  R  D  W  W  N  A  E  S  V  T  D  F  W  Q  N  W  N  I  P  V  172

```

21/33

FIG. 6A

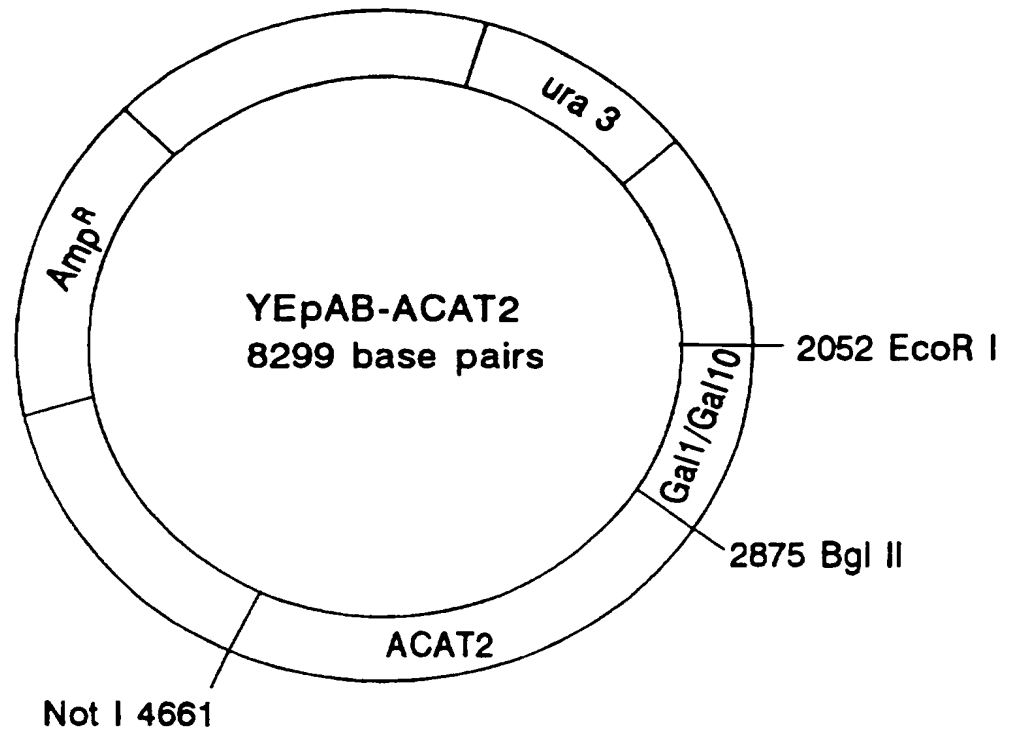
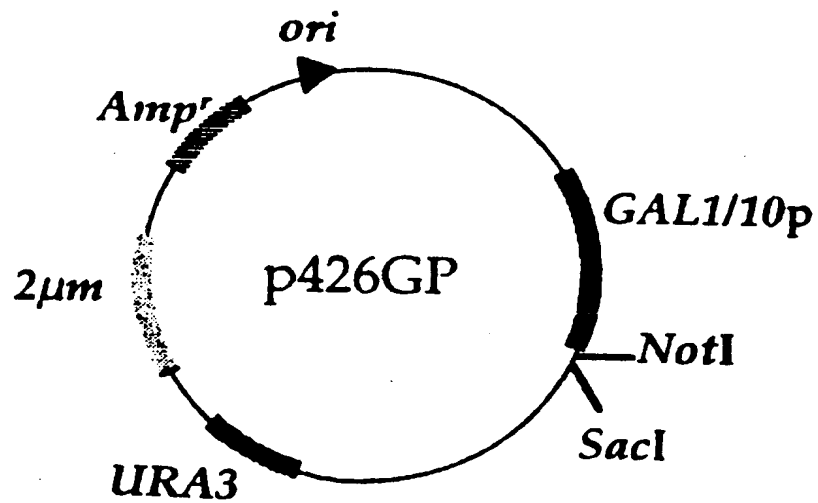
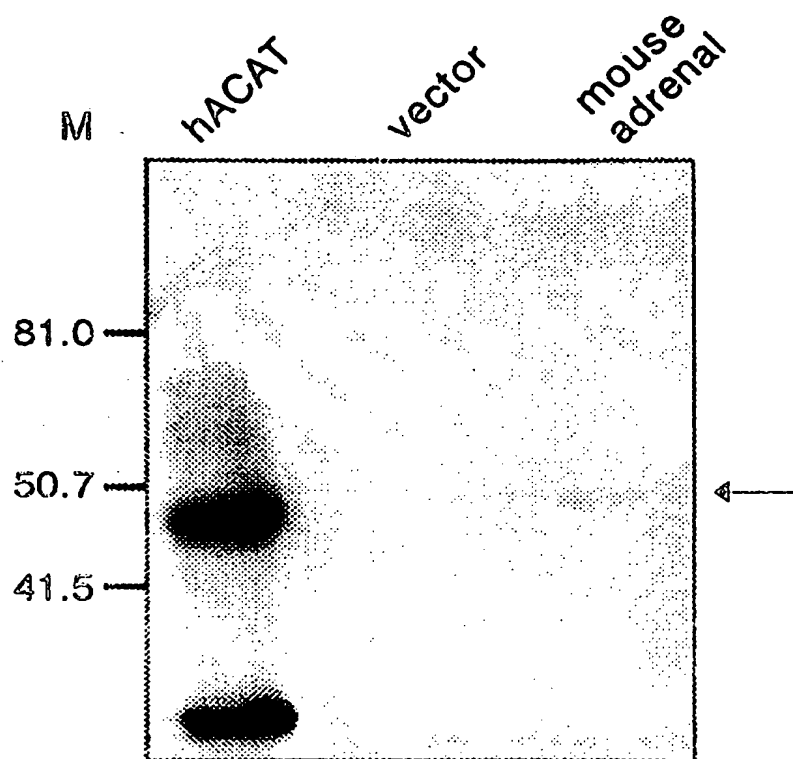


FIG. 6B



22/33

FIG. 6C

 α hACAT, DM10

SUBSTITUTE SHEET (RULE 26)

23/33

FIG. 7B

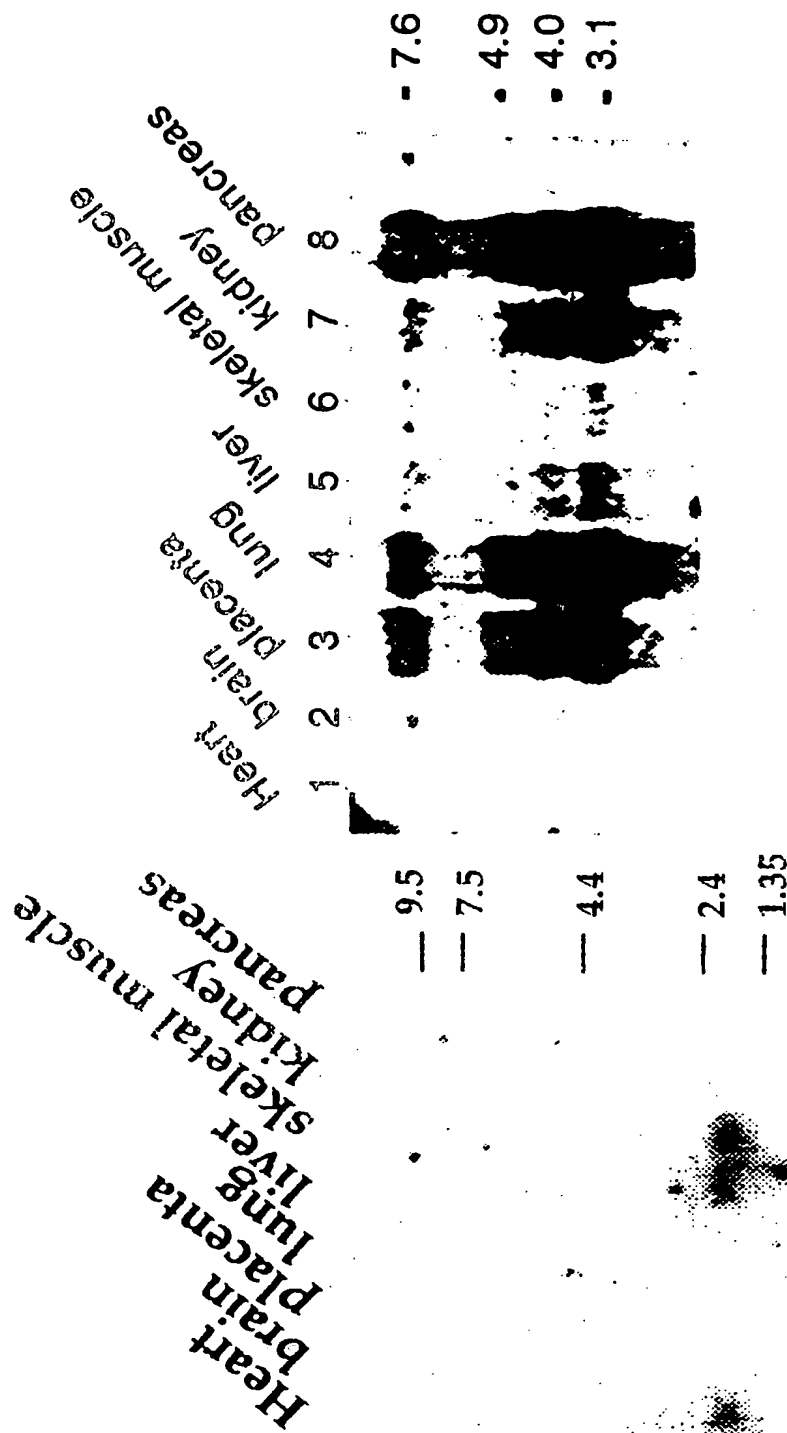


FIG. 7A

24/33

FIG. 8A

FIG. 8B

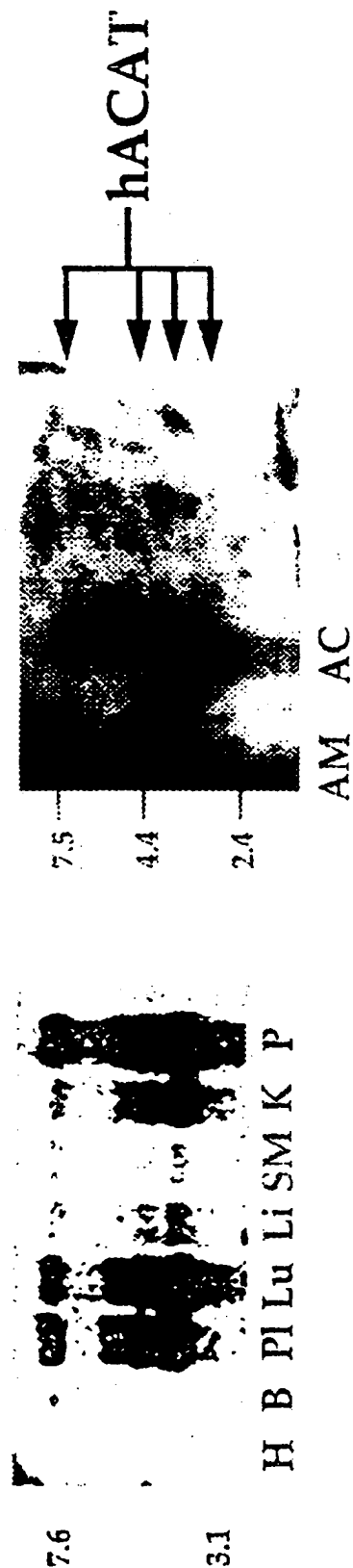
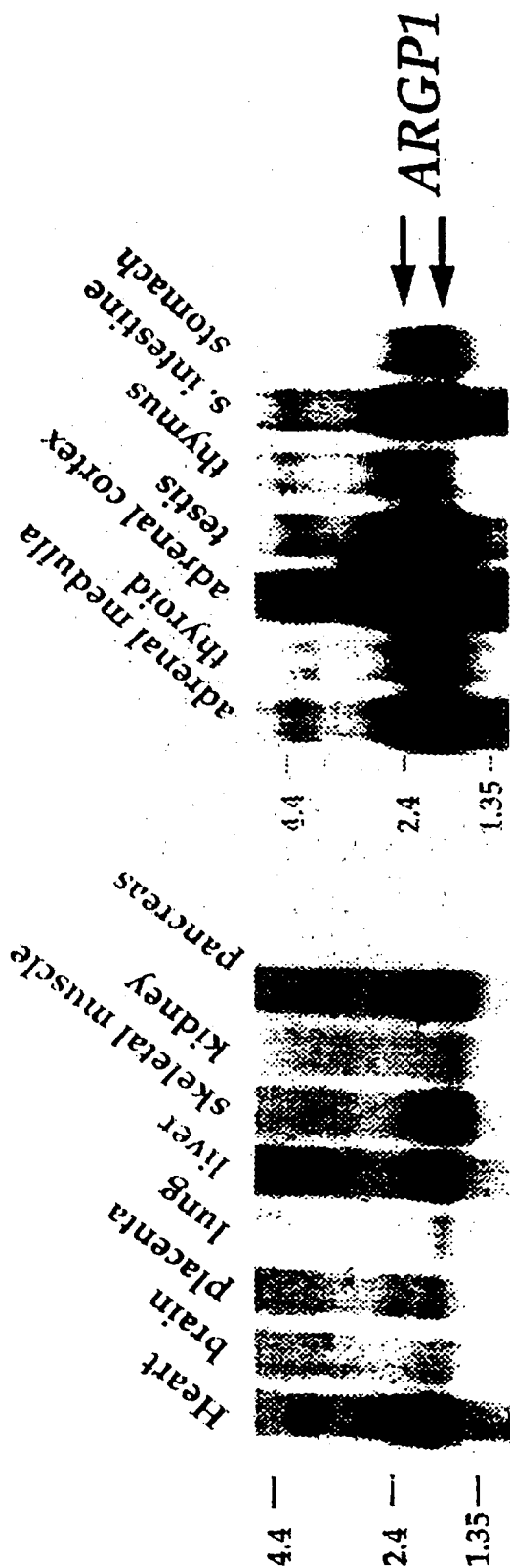
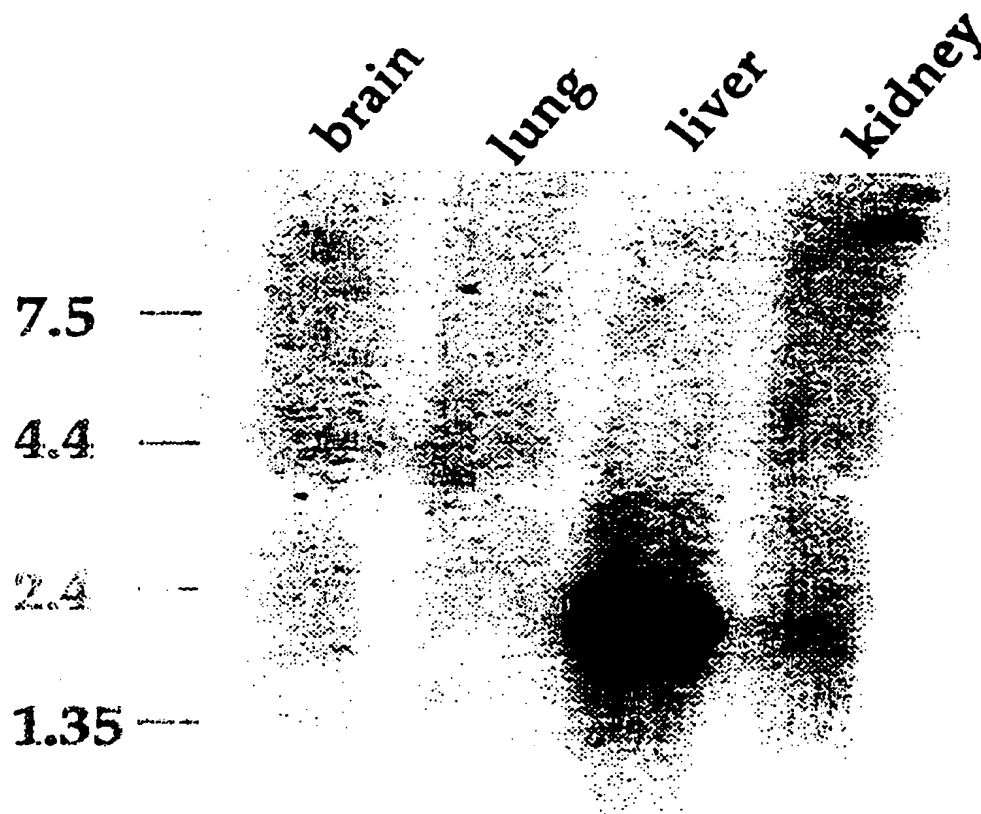


FIG. 8C

FIG. 8D

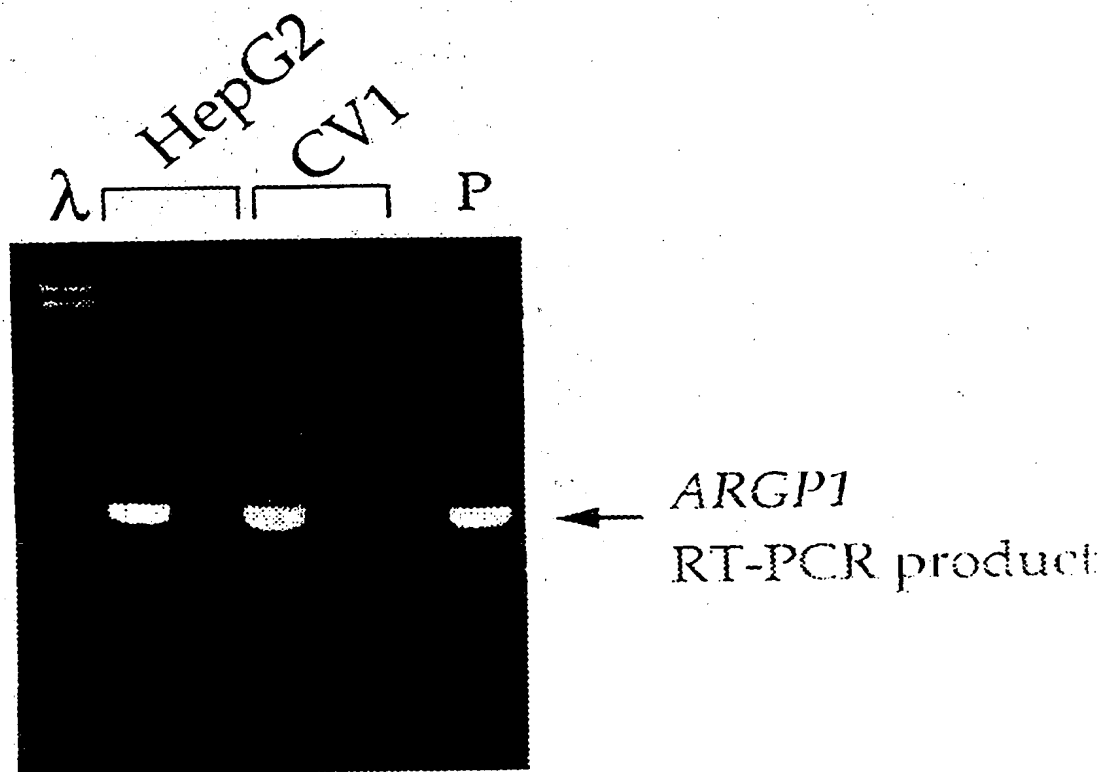
25/33

FIG. 9



26/33

FIG. 10



SUBSTITUTE SHEET (RULE 26)

27/33

FIG. 11

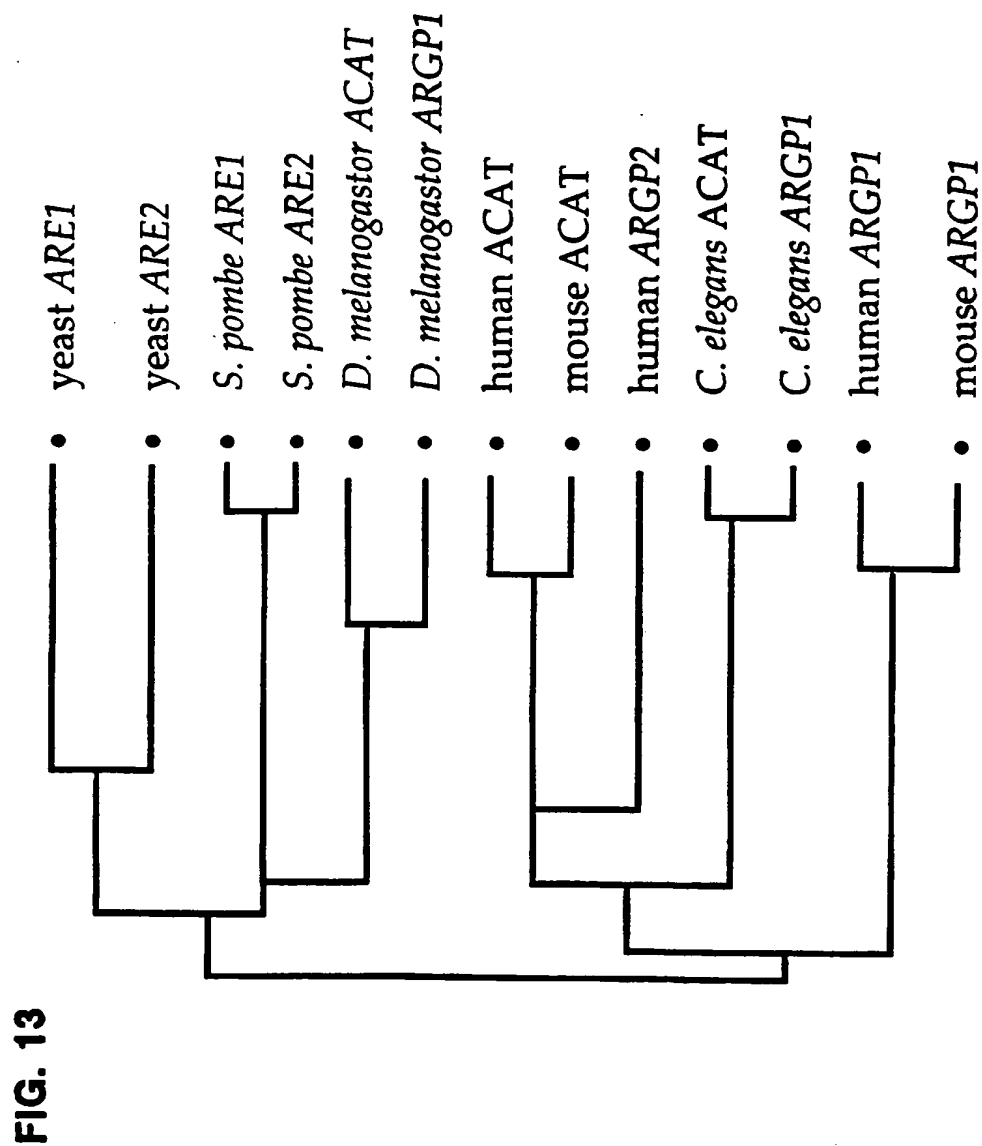
ARGP1	1MSNYRGILNWCVVMLI	15
	:. :.:.	
hACAT	101	EGEKNNHRAKDLRAPPEQGKIFIARRSLDELLEVDHIRTYYHMFIALLI	150
ARGP1	16	LSNARLFLENLIKYGILVDPIQVVSFLKDPHSWPAPCLVIAANVFAVAA	65
		. :. : . :. : . : . . . :. : . . .:	
hACAT	151	LFILSTLVVDYIDEGRLVLEFSLLSYAFGKFPTVVWTTWIMFLSTFSVPY	200
ARGP1	66	FQVEKRLAVGALTEQA.....GLLLHVANLATILCFPAAVVLLVESIT	108
		:.:. . . :. : : : :. : : . .:. : . . :	
hACAT	201	FLFQHWRTGYSKSSHPLIRSLFHGFLFMIFQIG.VLGFGPTYVVLAYTLP	249
ARGP1	109	PVGSLALMAHTILFLKPFYSYRDVNSWCRRARAKAASAGKKASSAAAPHT	158
		:.:. :. :. :. : : : : . :. :. : . :. :	
hACAT	250	PASRFIIIFEQIRFVMA.....HSFVRENVPRVLNSAKEKSSTVPIPT	293
ARGP1	159	VSYPDNLTYRDLYYFLFAPTLCYELNFPSPRIRKRFLRLRRILEMLFFTQ	208
		. :. : : : . :. :. : .	
hACAT	294	VN.....QYLYFLFAPTLIYRDSYPRNPTVRWGYVAMK.FAQVFGCF	334
ARGP1	209	LQVGLIQQWMVPTIQNSMKPFKDMDISRIERLLKLAV.....PNHLIWL	253
		: . :. :. :. : :. :. : . . :.	
hACAT	335	FYVYYIFERLCAPL.....FRNIKQEPFSARVLVLCVFNSILPGVLILF	378
ARGP1	254	IFFYWLFHSCNNAVAELMQFGDREFYRDWWNSESVTYFWQWNIPVHKWC	303
		: : : : :. : . :. : :	
hACAT	379	LTFFAFLHCWLNFAEMLRFGDRMFYKDWWNSTSYSNYYRTWNVVVDWL	428
ARGP1	304	IRHFYKPMRLRGSSKWM..ARTGVFLASAFFHEYLVSVPLR.....MFRL	346
		. . : : :. : . : :	
hACAT	429	YYYAYKDFLWFFSKRFKSAAMLAVFAVSAVVHEYALAVCLSFFYPVLFVL	478
ARGP1	347	WAFMGMAQIPLAWFVGRFFQGNYGNAAVWLSLIIGQPIAVLMYVHDYYV	396
		: :. :. : . . . : : :. : : :. :	
hACAT	479	FMFFGM....AFNFIVNDSRKKPIWNVLMWTSFLGNGVLLCFYSQEWYA	524
ARGP1	397	LNIEAPAAEA.....	406
		. . : . .	
hACAT	525	RRHCPLKNPTFLDYVRPRSWTCRYVF	550

28/33

FIG. 12

ARGP2	1M.FTPALGCVFYACFIVARFCVPVFA	24
hACAT	301	LFAPTLIYRDSYPRNPTVRWGYVAMKFAQVFGCFYVYIFERLCAPLFR	350
ARGP2	25	NMSREPFSTRALVFFILHATLPGIFMLLLIFFAFLHCWLNFAEMLRFGD	74
hACAT	351	NIKQEPFSARVLVLCVENSILPGVLILFLTFFAFLHCWLNFAEMLRFGD	400
ARGP2	75	RMFYRDWNNSTSFNSYRTWNVVHDWLYSYVYQDGLRLLGARARGVAML	124
hACAT	401	RMFYKDWNSTSYNSYRTWNVVHDWLYYAYKDFLWFFSKRFSAAAML	450
ARGP2	125	GVFLVSAVAHEYIFCFVLGFFYPVMLILFLVIEGMLNFMHMQRTGPWN	174
hACAT	451	AVFAVSADVHEYALAVCLSFYPVLFVLFEMFFGMAFNIVNDSRKKPIWN	500
ARGP2	175	VLMWTMLFLGQGIQVSLYCQEWYARRHCPLPQATFWGLVTPRSWSCHT..	222
hACAT	501	VLMWTSFLGNGVLLCFYSQEWYARRHCPLKNPTFELDYVRPRSWTCRYVF	550

29/33



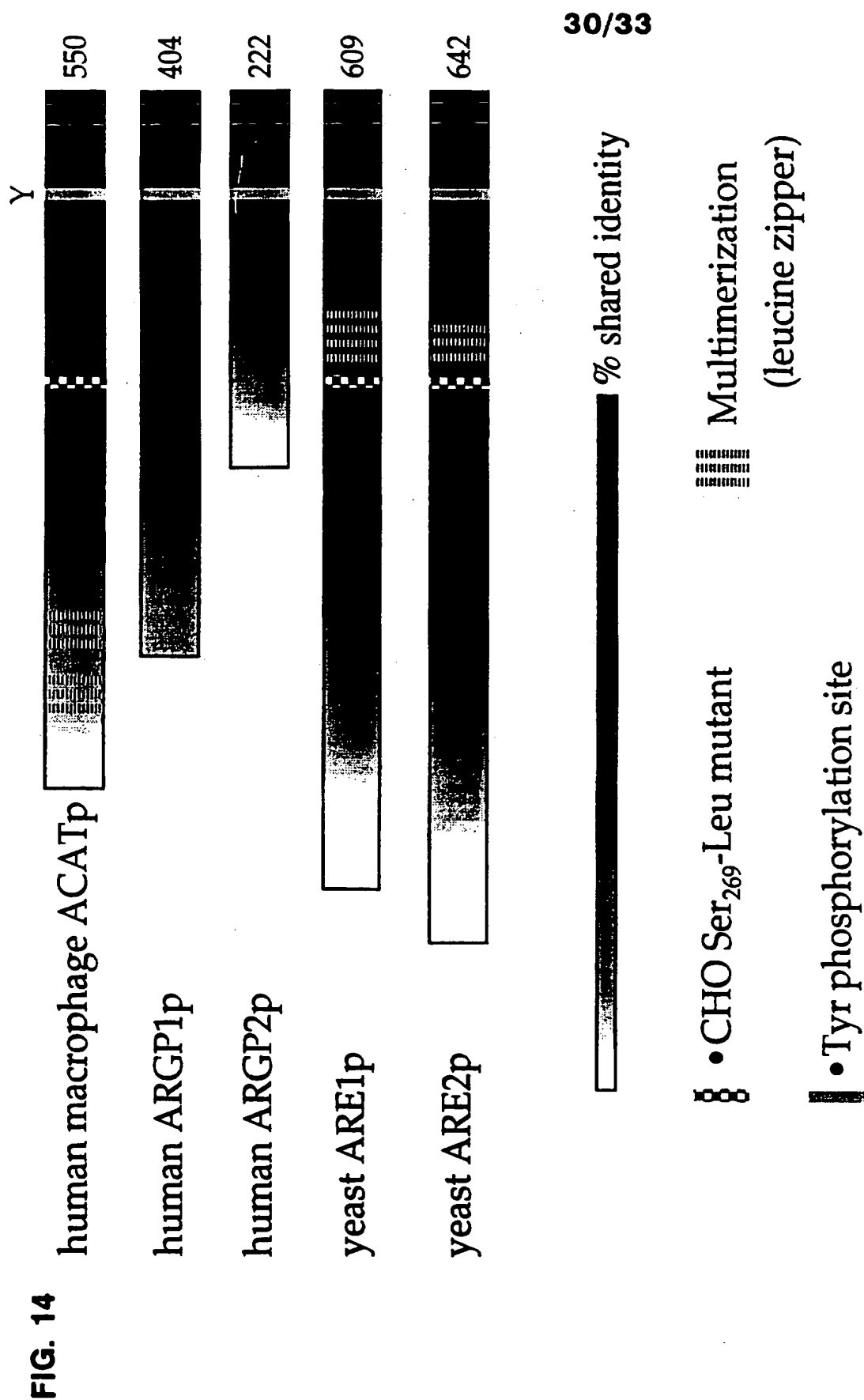


FIG. 15A

1	ATG	AGC	AAC	TAC	CGT	GCC	ATC	CTG	AAC	TGG	TGT	GTG	GTG	ATG	CTG	ATC	TTC	AGC	AAT	GCC	CGG	TTA	TTT	CTG	GAG	75
1	M	S	N	Y	R	G	I	L	N	W	C	V	V	M	L	I	L	S	N	A	R	L	F	L	E	25
76	AAC	CTC	ATC	AAG	TAT	GCC	ATC	CTG	GTG	GAC	CCC	ATC	CAG	GTG	GTT	TCT	CTG	TTC	CTG	AAG	GAT	CCC	CAT	AGC	TGG	150
26	N	L	I	K	Y	G	I	L	V	D	P	I	Q	V	V	S	L	F	L	K	D	P	H	S	W	50
151	CCC	GCC	CCA	TGC	CTG	GTT	ATT	GCG	GCC	AAT	GTC	TTT	GCT	GTG	GCT	GCA	TTC	CAG	GTT	GAG	AAG	CGC	CTG	GCG	GTG	225
51	P	A	P	C	L	V	I	A	A	N	V	F	A	V	A	A	F	Q	V	E	K	R	L	A	V	75
226	GGT	GCC	CTG	ACG	GAG	CAG	GCG	GGA	CTG	CTG	CTG	CAC	GTA	GCC	AAC	CTG	GCC	ACC	ATT	CTG	TGT	TTC	CCA	GCG	GCT	300
76	G	A	L	T	E	Q	A	G	L	L	L	H	V	A	N	L	A	T	I	L	C	F	P	A	A	100
301	GTG	GTC	TTA	CTG	GTT	GAG	TCT	ATC	ACT	CCA	GTG	GCG	TCC	CTG	CTG	GCG	CTG	ATG	GCG	CAC	ACC	ATC	CTC	TTC	CTC	375
101	V	V	L	L	V	E	S	I	T	P	V	G	S	L	L	A	L	M	A	H	T	I	L	F	L	125
376	AAG	CCC	TTC	TCC	TAC	CGC	GAC	GTC	AAC	TCA	TGG	TGC	CGC	AGG	GCC	AGG	GCC	AAG	GCT	GCC	TCT	GCA	GCG	AAG	AAG	450
126	K	P	F	S	Y	R	D	V	N	S	W	C	R	A	R	A	K	A	A	S	A	G	K	K	K	150
451	GCC	AGC	AGT	GCT	GCT	GCC	CCG	CAC	ACC	GTG	AGC	TAC	CCG	GAC	AAT	CTG	ACC	TAC	CGC	GAT	CTC	TAC	TAC	TTC	CTC	525
151	A	S	S	A	A	A	P	H	T	V	S	Y	P	D	N	L	T	Y	R	D	L	Y	Y	F	L	175
526	TTC	GCC	CCC	ACC	TTG	TGC	TAC	GAG	CTC	AAC	TTT	CCC	CGC	TCT	CCC	CGC	ATC	CGG	AAG	CGC	TTT	CTG	CTG	CGA	CGG	600
176	F	A	P	T	L	C	Y	E	L	N	F	P	R	S	P	R	I	R	K	R	F	L	L	R	R	200
601	ATC	CTT	GAG	ATG	CTG	TTC	TTC	ACC	CAG	CTC	CAG	GTG	GCG	CTG	ATC	CAG	CAG	TGG	ATG	GTG	CCC	ACC	ATC	CAG	AAC	675
201	I	L	E	M	L	F	F	T	Q	L	Q	V	G	L	I	Q	Q	W	M	V	P	T	I	Q	N	225
676	TCC	ATG	AAG	CCC	TTC	AAG	GAC	ATG	GAC	TAC	TCA	CGC	ATC	ATC	GAG	CGC	CTC	CTG	AAG	CTG	GCG	GTG	CCC	AAT	CAC	750
226	S	H	K	P	F	K	D	M	D	Y	S	R	I	I	E	R	L	L	K	L	A	V	P	N	H	250

31/33

32/33

FIG. 15B

751 CTC ATC TGG CTC ATC TTC TTC TAC TGG CTC TTC CAC TCC TGC CTG AAT GCC GTG GCT GAG CTC ATG CAG TTT GGA 825
 251 L I W L I F Y W L F H S C L N A V A E L M Q F G 275
 826 GAC CGG GAG TTC TAC CGG GAC TGG TGG AAC TCC GAG TCT GTC ACC TAC TTC TGG CAG AAC TGG AAC ATC CCT GTG 900
 276 D R E F Y R D W W N S E S V T Y F W Q N W N I P V 300
 901 CAC AAG TGG TGC ATC AGA CAC TTC TAC AAG CCC ATG CTT CGA CGG GGC AGC AGC AAG TGG ATG GCC AGG ACA GGG 975
 301 H K W C I R H F Y K P M L R R G S S K W M A R T G 325
 976 GTG TTC CTG GCC TGG GCT TTC CAC GAG TAC CTG GTG AGC GTC CCT CTG CGA ATG TTC CGC CTC TGG GCT TTC 1050
 326 V F L A S A F F H E Y L V S V P L R M F R L W A F 350
 1051 ACG GGC ATG ATG GCT CAG ATC CCA CTG GCC TGG TTC GTG GGC CGC TTT TTC CAG GGC AAC TAT GGC AAC GCA GCT 1125
 351 T G M A Q I P L A W F V G R F F Q G N Y G N A A 375
 1126 GTG TGG CTG ATC ATC ATC GGA CAG CCA ATA GCC GTC CTC ATG TAC GTC CAC GAC TAC TAC GTG CTC AAC TAT 1200
 376 V W L S L I I G Q P I A V L M Y V H D Y Y V L N Y 400
 1201 GAG GCC CCA CGG GCA GAG GCC TGA gctgcacctgaggccctggcttctcactgccacctcaaacccgctgccagagccacctctcctcta 1292
 401 E A P A A E A . 408
 1293 ggcctcgagtgctgggctggctgcacagcatcctcctctggtcccgaggagccctctctgccctatggggctctgtctgcacccctcaggga 1392
 1393 tggcgacagcaggccagacacagtgatgccagctgggagtgctgtgacccctgccccgggtccgaggggtgtcaataaagtgtgtccagtgaggagaaa 1492
 1493 aaaaaaaaaaaaaattctgcggccgc 1521

FIG. 16

1 ATG TTT ACC CCA GCC CTG GGA TGT GTG TTC TAT GCC TGC TTC ATC GTG GCC CGC TTC TGT GTT CCT GTC TTT GCC 75
 1 M F T P A L G C V F Y A C F I V A R F C V P V F A 25
 76 AAC ATG AGC CGA GAG CCC TTC AGC ACC CGT GCC TTG GTG TTT TTT ATC CTG CAT GCC ACG TTG CCA GGC ATT TTC 150
 26 N M S R E P F S T R A L V F F I L H A T L P G I F 50
 151 ATG CTG CTG ATC CTC TTT GCC TTC CTG TGC TGG CTC AAC GCC TTT GCC GAG ATG CTA CGA TTT GGA GAC 225
 51 M L L L I F A F L H C W L N A F A E M L R F G D 75
 226 AGG ATG TTC TAC CGG GAC TGG TGG AAC TCA ACG TCC TTC TCC AAC TAC TAC CGC ACT TGG AAC GTG GTG CAT 300
 76 R M F Y R D W W N S T S F S N Y Y R T W N V V V H 100
 301 GAC TGG CTG TAC AGC TAC GTG TAT CAG GAT GGG CTG CGG CTC CTT GGT GCC CGG GGC CGA GGG GTA GCC ATG CTG 375
 101 D W L Y S Y V Y Q D G L R L G A R A R G V A M L 125
 376 GGT GTG TTT CTG GTC TCC GCA GTG GCC CAC GAG TAT ATC TTC TGC TTC GTG GGG TTC TTC TAT CCC GTC ATG 450
 126 G V F L V S A V A H E Y I F C F V L G F F Y P V M 150
 451 CTG ATA CTC TTT GTC ATT GAA GGA ATG TTG AAC TTC ATG CAT GAC CAG CGC ACC GGC CCG GCA TGG AAC 525
 151 L I L F L V I E G M L N F M H D Q R T G P A W N 175
 526 GTG CTG ATG TGG ACC ATG CTG TTT CTA GGC CAG GGA ATC CAG GTC AGC CTG TAC TGC CAG GAG TGG TAC GCA CGG 600
 176 V L M W T M L F L G Q G I Q V S L Y C Q E W Y A R 200
 601 CGG CAC TGC CCC TTA CCC CAG GCA ACT TTC TGG GGG CTG GTG ACA CCT CGA TCT TGG TCC TGC CAT ACC TAG aggt 676
 201 R H C P L P Q A T F W G L V T P R S W S C H T 224
 677 cgggacagacgactgcccagacaccaccaaagtctctgcctgcaaaacctggggaccaggactccctgtctgcattccccaaattggctctga 776
 777 gtcgaggcaacctgcacaaagacccccaccaccaaaggaatgtgcaaggactgagatctgcagactgtgtggtaactgatcacagacctcagcatgggggtg 876
 877 accagggtgactcttcaatccctatccccatgggctgggtacaggatatccctcctaccatgactgtcttagggagacttggggtacaccttgatttg 976
 977 atgaatgtgggggaactcagaggaactggggccaccaagggttggaagggtttggttcttgactttgtattccttccaataacagcaataaactttgtct 1076
 1077 cccttt 1082

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/09460**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(6) :C07H 21/04; C12N 15/11; C12P 21/02

US CL :536/23.2, 23.5; 435/70.1, 320.1, 325

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.2, 23.5; 435/70.1, 320.1, 325

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, MEDLINE, BIOSIS, CA, DERWENT

search terms: acylcoenzyme A: cholesterol acyltransferase

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X --- Y	Expressed Sequence Tag (EST) database Accession Numbers R99213 (13 September 1995), R99214 (13 September 1995), H27149 (12 July 1995), N76754 (02 April 1996), H45923 (31 July 1995), H45924 (31 July 1995), H24791 (07 July 1995), Z39933 (21 September 1995), T79408 (15 March 1995), R48474 (18 May 1995), T79494 (15 March 1995), R48475 (18 May 1995), M79086 (26 May 1992), and W10786 (26 April 1996).	1-12 and 33 ----- 13-32
Y	US 5,484,727 A (CHANG et al.) 16 January 1996, cols. 1-5.	13-32

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier document published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

19 AUGUST 1997

Date of mailing of the international search report

29 AUG 1997

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

KENNETH R. BORLICK

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/09460

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-33

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/09460

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-33, drawn to acylcoenzyme A: cholesterol acyltransferase II and III nucleic acids and enzymes, and a first method of use comprising recombinant expression.

Group II, claim(s) 34-39, drawn to a second method of use comprising detecting a defect in esterification of sterol in a subject.

Group III, claim(s) 40, drawn to third method of use comprising treating a subject having an imbalance in sterol levels.

Group IV, claims 41-44, drawn to a fourth method of use comprising inhibiting wildtype acylcoenzyme A: cholesterol acyltransferase II or III in a subject.

Group V, claim 45, drawn to a fifth method of use comprising identifying a chemical compound which is capable of inhibiting acylcoenzyme A: cholesterol acyltransferase II or III in a subject.

Group VI, claims 46-48, drawn to an undefined pharmaceutical compound and methods of its use in treating a subject.

Group VII, claim 49, drawn to a transgenic, nonhuman mammal.

The inventions listed as Groups I-VII do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: PCT Rule 13.2 only provides for grouping together of a first product and a first method of use, rather than for multiple uses. Further, the compound and uses of group VI and the transgenic mammal of group VII lack the same or corresponding special technical feature as the other groups, which relate specifically to the acylcoenzyme A:cholesterol acyltransferase II or III sequence.

